5

Stochastic Differential Equations

We shall now solve the stochastic differential equation of section 1.1, which served as the motivation for the stochastic integration theory developed so far.

5.1 Introduction

The stochastic differential equation (1.1.9) reads[1]

$$X_t = C_t + \int_0^t F_\eta[X] \, dZ^\eta \ . \tag{5.1.1}$$

The previous chapters were devoted to giving a meaning to the dZ^η-integrals and to providing the tools to handle them. As it stands, though, the equation above still does not make sense; for the solution, if any, will be a right-continuous adapted process – but then so will the $F_\eta[X]$, and we cannot in general integrate right-continuous integrands. What does make sense in general is the equation

$$X_t = C_t + \int_0^t F_\eta[X]_{s-} \, dZ_s^\eta \tag{5.1.2}$$

or, equivalently, $$X = C + F_\eta[X]_{\leftharpoondown} * Z^\eta = C + \boldsymbol{F}[X]_{\leftharpoondown} * \boldsymbol{Z} \tag{5.1.3}$$

in the notation of definition 3.7.6, and with $\boldsymbol{F}[X]_{\leftharpoondown}$ denoting the left-continuous version of the matrix

$$\boldsymbol{F}[X] = \big(F_\eta[X]\big)_{\eta=1\ldots d} = \big(F_\eta^\nu[X]\big)_{\eta=1\ldots d}^{\nu=1\ldots n} \ .$$

In (5.1.3) now the integrands are left-continuous, and therefore previsible, and thus are integrable if not too big (theorem 3.7.17).

The intuition behind equation (5.1.2) is that $X_t = (X_t^\nu)^{\nu=1\ldots n}$ is a vector in $\mathbb{R}^n$ representing the state at time t of some system whose evolution is

[1] Einstein's convention is adopted throughout; it implies summation over the same indices in opposite positions, for instance, the η in (5.1.1).

driven by a collection $\boldsymbol{Z} = \{Z^\eta : 1 \le \eta \le d\}$ of scalar integrators. The F_η are the coupling coefficients, which describe the effect of the background noises Z^η on the change of X. $C = (C^\nu)^{\nu=1\ldots n}$ is the initial condition. Z^1 is typically time, so that $F_1[X]_{t-}\, dZ^1_t = F_1[X]_{t-}\, dt$ represents the systematic drift of the system.

First Assumptions on the Data and Definition of Solution

The ingredients X, C, F_η, Z^η of equation (5.1.2) are for now as general as possible, just so the equation makes sense. It is well to state their nature precisely:

(i) $Z^1, \ldots, Z^d$ are L^0-integrators. This is the very least one must assume to give meaning to the integrals in (5.1.2). It is no restriction to assume that $Z^\eta_0 = 0$, and we shall in fact do so. Namely, by convention $F_\eta[X]_{0-} = 0$, so that Z^η and $Z^\eta - Z^\eta_0$ has the same effect in driving X.

*(ii) The **coupling coefficients** F_η are random vector fields.* This means that each of them associates with every n-vector of processes[2] $X \in \mathfrak{D}^n$ another n-vector $F_\eta[X] \in \mathfrak{D}^n$. Most frequently they are ***markovian***; that is to say, there are ordinary ***vector fields*** $f_\eta : \mathbb{R}^n \to \mathbb{R}^n$ such that F_η is simply the composition with f_η: $F_\eta[X]_t = f_\eta(X_t) = f_\eta \circ X_t$. It is, however, not necessary, and indeed would be insufficient for the stability theory, to assume that the correspondences $X \mapsto F_\eta[X]$ are always markovian. Other coefficients arising naturally are of the form $X \mapsto A_\eta \cdot X$, where A_η is a bounded matrix-valued process[2] in $\mathfrak{D}^{n\times n}$. Genuinely non-markovian coefficients also arise in context with approximation schemes (see equation (5.4.29) on page 323.)

Below, Lipschitz conditions will be placed on $\boldsymbol{F}$ that will ensure automatically that the F_η are ***non-anticipating*** in the sense that at all stopping times T

$$F_\eta[X]_T = F_\eta[X^T]_T \,. \tag{5.1.4}$$

To paraphrase: "at any time T the value of the coupling coefficient is determined by the history of its argument up to that time." It is sometimes convenient to require that $F_\eta[0] = 0$ for $1 \le \eta \le d$. This is no loss of generality. Namely, X solves equation (5.1.3):

$$X = C + F_\eta[X]_{.-} {*} Z^\eta = C + \boldsymbol{F}[X]_{.-} {*} \boldsymbol{Z} \tag{5.1.5}$$

iff it solves
$$X = {}^0C + {}^0F_\eta[X]_{.-} {*} Z^\eta = {}^0C + {}^0\boldsymbol{F}[X]_{.-} {*} \boldsymbol{Z} \,, \tag{5.1.6}$$

where ${}^0C \overset{\text{def}}{=} C + \boldsymbol{F}[0]_{.-} {*} \boldsymbol{Z}$ is the adjusted initial condition and where ${}^0\boldsymbol{F}[X] \overset{\text{def}}{=} \boldsymbol{F}[X]_{.-} - \boldsymbol{F}[0]_{.-}$ is the adjusted coupling coefficient, which vanishes at $X = 0$. Note that ${}^0C$ will not, in general, be constant in time, even if C is.

[2] Recall that $\mathfrak{D}$ is the vector space of all càdlàg adapted processes and $\mathfrak{D}^n$ is its n-fold product, identified with the càdlàg $\mathbb{R}^n$-valued processes.

(iii) $C \in \mathfrak{D}^n$. In other words, C is an n-vector of adapted right-continuous processes with left limits. We shall refer to C as the ***initial condition***, despite the fact that it need not be constant in time. The reason for this generality is that in the stability theory of the system (5.1.2) time-dependent additive random inputs C appear automatically – see equation (5.2.32) on page 293. Also, in the form (5.1.6) the random input ${}^0C$ is time-dependent even if the C in the original equation is not.

(iv) ***X solves the stochastic differential equation on the stochastic interval*** $[\![0,T]\!]$ if the stopped process X^T belongs to $\mathfrak{D}^n$ and[3]

$$X^T = C^T + \boldsymbol{F}[X]_{\cdot-} * \boldsymbol{Z}^T$$

or, equivalently, if

$$X^T = {}^0C^T + {}^0\boldsymbol{F}[X]_{\cdot-} * \boldsymbol{Z}^T \; ;$$

we also say that X ***solves the equation up to time*** T, or that X is a ***strong solution*** on $[\![0,T]\!]$. In view of theorem 3.7.17 on page 137, there is no question that the indefinite integrals on the right-hand side exist, at least in the sense L^0. Clearly, if X solves our equation on both $[\![0,T_1]\!]$ and $[\![0,T_2]\!]$, then it also solves it on the union $[\![0,T_1 \vee T_2]\!]$. The supremum of the (classes of the) stopping times T so that X solves our equation on $[\![0,T]\!]$ is called the ***lifetime*** of the solution X and is denoted by $\zeta[X]$. If $\zeta[X] = \infty$, then X is a ***(strong) global solution***.

As announced in chapter 1, stochastic differential equations will be solved using Picard's iterative scheme. There is an analog of Euler's method that rests on compactness and works when the coupling coefficients are merely continuous. But the solution exists only in a weak sense (section 5.5), and to extract uniqueness in the presence of some positivity is a formidable task (ibidem, [105]).

We next work out an elementary example. Both the results and most of the arguments explained here will be used in the stochastic case later on.

Example: The Ordinary Differential Equation (ODE)

Let us recall how Picard's scheme, which was sketched on pages 1 and 5, works in the deterministic case, when there is only one driving term, time. The stochastic differential equation (5.1.2) is handled along the same lines; in fact, we shall refer below, in the stochastic case, to the arguments described here without carrying them out again. A few slightly advanced classical results are developed here in detail, for use in the approximation of Stratonovich equations (see page 321 ff.).

The ordinary differential equation corresponding to (5.1.2) is

$$x_t = c_t + \int_0^t f(x_s)\, ds \; . \tag{5.1.7}$$

[3] If X is only defined on $[\![0,T]\!]$, then X^T shall mean that extension of X which is constant on $[\![T,\infty)\!)$.

To paraphrase: "time drives the state in the direction of the vector field f." To solve it one defines for every path $x_. : [0,\infty) \to \mathbb{R}^n$ the path $\mathfrak{u}[x_.]_.$ by

$$\mathfrak{u}[x_.]_t = c_t + \int_0^t f(x_s)\,ds\ , \qquad t \ge 0\ .$$

Equation (5.1.7) asks for a fixed point of the map $\mathfrak{u}$. Picard's method of finding it is to design a complete norm $\|\ \ \|$ on the space of paths with respect to which $\mathfrak{u}$ is ***strictly contractive***. This means that there is a $\gamma < 1$ such that for any two paths $x'_.\,, x_.$

$$\left\| \mathfrak{u}[x'_.] - \mathfrak{u}[x_.] \right\| \le \gamma \cdot \left\| x'_. - x_. \right\|\ .$$

Such a norm is called a ***Picard norm*** for $\mathfrak{u}$ or for (5.1.7), and the least γ satisfying the previous inequality is the ***contractivity modulus*** of $\mathfrak{u}$ for that norm. The map $\mathfrak{u}$ will then have a fixed point, solution of equation (5.1.7) – see page 275 for how this comes about.

The strict contractivity of $\mathfrak{u}$ usually derives from the assumption that the vector field f on $\mathbb{R}^n$ is ***Lipschitz***; that is to say, that there is a constant $L < \infty$, the ***Lipschitz constant*** of f, such that[4]

$$\left| f(x'_t) - f(x_t) \right| \le L \cdot \left| x'_t - x_t \right| \quad \forall\, t \ge 0\ . \tag{5.1.8}$$

For a path $x_.$ define as usual its maximal path $|x|^\star_.$ by $|x|^\star_t = \sup_{s \le t} |x_s|$, $t \ge 0$. Then inequality (5.1.8) has the consequence that for all t

$$\left| f(x'_.) - f(x_.) \right|^\star_t \le L \cdot \left| x'_. - x_. \right|^\star_t\ . \tag{5.1.9}$$

If this is satisfied, then

$$\|\, x_. \,\| = \|\, x_. \,\|_M \stackrel{\text{def}}{=} \sup_{t>0}\ e^{-Mt} \cdot |x|^\star_t = \sup_{t>0}\ e^{-Mt} \cdot |x|_t \tag{5.1.10}$$

is a clever choice of the Picard norm sought, just as long as M is chosen strictly greater than L. Namely, inequality (5.1.9) implies

$$\left| f(x') - f(x) \right|^\star_t \le Le^{Mt} \cdot e^{-Mt} \left| x' - x \right|^\star_t \le Le^{Mt} \cdot \left\| x'_. - x_. \right\|_M\ . \tag{5.1.11}$$

Therefore, multiplying the ensuing inequality

$$\begin{aligned} \left| \mathfrak{u}[x'_.]_t - \mathfrak{u}[x_.]_t \right| &\le \left| \int_0^t f(x'_s) - f(x_s)\,ds \right| \le \int_0^t \left| f(x') - f(x) \right|^\star_s ds \\ &\le L \cdot \left\| x'_. - x_. \right\|_M \cdot \int_0^t e^{Ms}\,ds \le \frac{L}{M} \cdot \left\| x'_. - x_. \right\|_M \cdot e^{Mt} \end{aligned}$$

by e^{-Mt} and taking the supremum over t results in

$$\left\| \mathfrak{u}[x'_.] - \mathfrak{u}[x_.] \right\|_M \le \gamma \cdot \left\| x'_. - x_. \right\|_M\ , \tag{5.1.12}$$

[4] $|\ \ |$ denotes the absolute value on $\mathbb{R}$ and also any of the usual and equivalent ℓ^p-norms $|\ \ |_p$ on $\mathbb{R}^k$, $\mathbb{R}^n$, $\mathbb{R}^{k\times n}$, etc., whenever p does not matter.

with $\gamma \overset{\text{def}}{=} L/M < 1$. Thus $\mathfrak{u}$ is indeed strictly contractive for $\| \ \|_M$.

The strict contractivity implies that $\mathfrak{u}$ has a unique fixed point. Let us review how this comes about. One picks an arbitrary starting path $x_.^{(0)}$, for instance $x_.^{(0)} \equiv 0$, and defines the Picard iterates by $x_.^{(n+1)} \overset{\text{def}}{=} \mathfrak{u}[x_.^{(n)}]_.$, $n = 0, 1, \ldots$. A simple induction on inequality (5.1.12) yields

$$\left\| x_.^{(n+1)} - x_.^{(n)} \right\|_M \le \gamma^n \cdot \left\| \mathfrak{u}[x_.^{(0)}] - x_.^{(0)} \right\|_M$$

and
$$\sum_{n=1}^{\infty} \left\| x_.^{(n+1)} - x_.^{(n)} \right\|_M \le \frac{\gamma}{1-\gamma} \cdot \left\| \mathfrak{u}[x_.^{(0)}] - x_.^{(0)} \right\|_M . \tag{5.1.13}$$

Provided that
$$\left\| \mathfrak{u}[x_.^{(0)}] - x_.^{(0)} \right\|_M < \infty , \tag{5.1.14}$$

the collapsing sum
$$x_. \overset{\text{def}}{=} x_.^{(1)} + \sum_{n=1}^{\infty} \Big(x_.^{(n+1)} - x_.^{(n)} \Big) = \lim_{n\to\infty} x_.^{(n)} \tag{5.1.15}$$

converges in the Banach space $\mathfrak{s}_M$ of paths $y_. : [0, \infty) \to \mathbb{R}^n$ that have $\|y_.\|_M < \infty$. Since $\mathfrak{u}[x_.] = \lim_n \mathfrak{u}[x^{(n)}] = \lim_n x^{(n+1)} = x$, the limit $x_.$ is a fixed point of $\mathfrak{u}$. If $x'_.$ is any other fixed point in $\mathfrak{s}_M$, then $\|x'_. - x_.\| \le \|\mathfrak{u}[x'_.] - \mathfrak{u}[x_.]\| \le \gamma \cdot \|x'_. - x_.\|$ implies that $\|x'_. - x_.\| = 0$: inside $\mathfrak{s}_M$, $x_.$ is the only solution of equation (5.1.7). There is a priori the possibility that there exists another solution in some space larger than $\mathfrak{s}_M$; generally some other but related reasoning is needed to rule this out.

Exercise 5.1.1 The norm $\| \ \|$ used above is by no means the only one that does the job. Setting instead $\|x_.\| \overset{\text{def}}{=} \int_0^\infty |x|_t^\star \cdot Me^{-Mt}\,dt$, with $M > L$, defines a complete norm on continuous paths, and $\mathfrak{u}$ is strictly contractive for it.

Let us discuss six consequences of the argument above — they concern only the action of $\mathfrak{u}$ on the Banach space $\mathfrak{s}_M$ and can be used literally or minutely modified in the general stochastic case later on.

5.1.2 General Coupling Coefficients To show the strict contractivity of $\mathfrak{u}$, only the consequence (5.1.9) of the Lipschitz condition (5.1.8) was used. Suppose that f is a map that associates with every path $x_.$ another one, but not necessarily by the simple expedient of evaluating a vector field at the values x_t. For instance, $f(x_.)_.$ could be the path $t \mapsto \phi(t, x_t)$, where ϕ is a measurable function on $\mathbb{R}_+ \times \mathbb{R}^n$ with values in $\mathbb{R}^n$, or it could be convolution with a fixed function, or even the composition of such maps. As long as inequality (5.1.9) is satisfied and $\mathfrak{u}[0]$ belongs to $\mathfrak{s}_M$, our arguments all apply and produce a unique solution in $\mathfrak{s}_M$.

5.1.3 The Range of $\mathfrak{u}$ Inequality (5.1.14) states that $\mathfrak{u}$ maps at least one – and then every – element of $\mathfrak{s}_M$ into $\mathfrak{s}_M$. This is another requirement on the system equation (5.1.7). In the present simple sure case it means that $\|c_. + \int_0^{\bullet} f(0)\,ds\|_M < \infty$ and is satisfied if $c_. \in \mathfrak{s}_M$ and $f(0)_. \in \mathfrak{s}_M$, since $\|\int_0^{\bullet} f(0)_s\,ds\|_M \le \|f(0)_.\|_M / M$.

5.1.4 Growth Control The arguments in (5.1.12)–(5.1.15) produce an a priori estimate on the growth of the solution $x_{\bullet}$ in terms of the initial condition and $\mathfrak{u}[0]$. Namely, if the choice $x^{(0)}_{\bullet} = 0$ is made, then equation (5.1.15) in conjunction with inequality (5.1.13) gives

$$\|x_{\bullet}\|_M \le \left\|x^{(1)}_{\bullet}\right\|_M + \frac{\gamma}{1-\gamma}\cdot\left\|x^{(1)}_{\bullet}\right\|_M = \frac{1}{1-\gamma}\cdot\left\|c_{\bullet}+\int_0^{\bullet} f(0)_s\,ds\right\|_M . \tag{5.1.16}$$

The very structure of the norm $\|\ \|_M$ shows that $|x_t|$ grows at most exponentially with time t.

5.1.5 Speed of Convergence The choice $x^{(0)}_{\bullet} \equiv 0$ for the zeroth iterate is popular but not always the most cunning. Namely, equation (5.1.15) in conjunction with inequality (5.1.13) also gives

$$\left\|x_{\bullet} - x^{(1)}_{\bullet}\right\|_M \le \frac{\gamma}{1-\gamma}\cdot\left\|x^{(1)}_{\bullet} - x^{(0)}_{\bullet}\right\|_M$$

and
$$\left\|x_{\bullet} - x^{(0)}_{\bullet}\right\|_M \le \frac{1}{1-\gamma}\cdot\left\|x^{(1)}_{\bullet} - x^{(0)}_{\bullet}\right\|_M .$$

We learn from this that if $x^{(0)}_{\bullet}$ and the first iterate $x^{(1)}_{\bullet}$ do not differ much, then both are already good approximations of the solution $x_{\bullet}$. This innocent remark can be parlayed into various schemes for the pathwise solution of a stochastic differential equation (section 5.4). For the choice $x^{(0)}_{\bullet} = c_{\bullet}$ the second line produces an estimate of the deviation of the solution from the initial condition:

$$\|x_{\bullet} - c_{\bullet}\|_M \le \frac{1}{1-\gamma}\cdot\left\|\int_0^{\bullet} f(c)_s\,ds\right\|_M \le \frac{1}{M(1-\gamma)}\cdot\|f(c)_{\bullet}\|_M . \tag{5.1.17}$$

5.1.6 Stability Suppose f' is a second vector field on $\mathbb{R}^n$ that has the same Lipschitz constant L as f, and $c'_{\bullet}$ is a second initial condition. If the corresponding map $\mathfrak{u}'$ maps 0 to $\mathfrak{s}_M$, then the differential equation $x'_t = c'_t + \int_0^t f'(x'_s)\,ds$ has a unique solution $x'_{\bullet}$ in $\mathfrak{s}_M$. The difference $\delta_{\bullet} \stackrel{\text{def}}{=} x'_{\bullet} - x_{\bullet}$ is easily seen to satisfy the differential equation

$$\delta_t = (c'-c)_t + \int_0^t g(\delta_s)\,ds ,$$

where $g : \delta \mapsto f'(\delta + x) - f(x)$ has Lipschitz constant L. Inequality (5.1.16) results in the estimates

$$\|x'_{\bullet} - x_{\bullet}\|_M \le \frac{1}{1-\gamma}\left\|(c'-c)_{\bullet} + \int_0^{\bullet} f'(x_s) - f(x_s)\,ds\right\|_M , \tag{5.1.18}$$

and
$$\|x_{\bullet} - x'_{\bullet}\|_M \le \frac{1}{1-\gamma}\left\|(c-c')_{\bullet} + \int_0^{\bullet} f(x'_s) - f'(x'_s)\,ds\right\|_M ,$$

reversing roles. Both exhibit neatly the dependence of the solution x on the ingredients c, f of the differential equation. It depends, in particular, Lipschitz-continuously on the initial condition c.

5.1.7 Differentiability in Parameters If initial condition and coupling coefficient of equation (5.1.7) depend differentiably on a parameter u that ranges over an open subset $U \subset \mathbb{R}^k$, then so does the solution. We sketch a proof of this, using the notation and terminology of definition A.2.45 on page 388. The arguments carry over to the stochastic case (section 5.3), and some of the results developed here will be used there.

Formally differentiating the equation $x[u]_t = c[u]_t + \int_0^t f(u, x[u]_s)\, ds$ gives

$$Dx[u]_t = \Big(Dc[u]_t + \int_0^t D_1 f(u, x[u]_s) ds\Big) + \int_0^t D_2 f(u, x[u]_s) \cdot Dx[u]_s ds \; . \quad (5.1.19)$$

This is a linear differential equation for an $n \times k$-matrix-valued path $Dx[u]_.$. It is a matter of a smidgen of bookkeeping to see that the remainder

$$Rx[v; u]_s \stackrel{\text{def}}{=} x[v]_s - x[u]_s - Dx[u]_s \cdot (v - u)$$

satisfies the linear differential equation

$$Rx[v; u]_t = \Big(Rc[v; u]_t + \int_0^t Rf\big(u, x[u]_s; v, x[v]_s\big)\, ds\Big) \qquad (5.1.20)$$

$$+ \int_0^t D_2 f\big(u, x[u]_s\big) \cdot Rx[v; u]_s\, ds \; .$$

At this point we should show that[5] $\big|Rx[v; u]_t\big| = o(|v - u|)$ as $v \to u$; but the much stronger conclusion

$$\|Rx[v; u]_.\|_M = o(|v - u|) \qquad (5.1.21)$$

seems to be in reach. Namely, if we can show that both

$$\|Rc[v; u]_.\| = o(|v - u|) \qquad (5.1.22)$$

and $$\|Rf(u, x[u]_.; v, x[v]_.)\| = o(|v - u|) \; ,$$

then (5.1.21) will follow immediately upon applying (5.1.16) to (5.1.20). Now (5.1.22) will hold if we simply require that $v \mapsto c[v]_.$, considered as a map from U to $\mathfrak{s}_M$, be uniformly differentiable; and this will for instance be the case if *the family* $\{c[\,\cdot\,]_t : t \geq 0\}$ *is uniformly equidifferentiable.*

Let us then require that $f : U \times \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable with bounded derivative $Df = (D_1 f, D_2 f)$. Then the common coupling coefficient $D_2 f(u, x)$ of (5.1.19) and (5.1.20) is bounded by $L \stackrel{\text{def}}{=} \sup_{u,x} \big\| D_2 f(u, x) \big\|_{\mathbb{R}^{k \times n} \to \mathbb{R}^n}$ (see exercise A.2.46 (iii)), and for every $M > L$ the solutions $x[u]_.$, $u \in U$, lie in a common ball of $\mathfrak{s}_M$. One hopes of course that the coupling coefficient

$$F : U \times \mathfrak{s}_M \to \mathfrak{s}_M \quad , \quad F : (u, x_.) \mapsto f(u, x_.)$$

[5] For $o(\,\cdot\,)$ and $O(\,\cdot\,)$ see definition A.2.44 on page 388.

is differentiable. Alas, it is not, in general. We leave it to the reader (i) to fashion a counterexample and (ii) to establish that F *is weakly differentiable* from $U \times \mathfrak{s}_M$ to $\mathfrak{s}_M$, *uniformly on every ball* [Hint: see example A.2.48 on page 389]. When this is done a first application of inequality (5.1.18) shows that $u \to x[u]_{\boldsymbol{.}}$ is Lipschitz from U to $\mathfrak{s}_M$, and a second one that $Rx[v;u]_t = o(|v-u|)$ on any ball of U: the solution $Dx[u]_{\boldsymbol{.}}$ of (5.1.19) really is the derivative of $v \mapsto x[v]_{\boldsymbol{.}}$ at u. This argument generalizes without too much ado to the stochastic case (see section 5.3 on page 298).

ODE: Flows and Actions

In the discussion of higher order approximation schemes for stochastic differential equations on page 321 ff. we need a few classical results concerning flows on $\mathbb{R}^n$ that are driven by different vector fields. They appear here in the form of propositions whose proofs are mostly left as exercises. We assume that the vector fields $f, g : \mathbb{R}^n \to \mathbb{R}^n$ appearing below are at least once differentiable with bounded and Lipschitz-continuous partial derivatives.

For every $x \in \mathbb{R}^n$ let $\xi_{\boldsymbol{.}} = \xi^f_{\boldsymbol{.}}(x) = \xi[x, \boldsymbol{\cdot}; f]$ denote the unique solution of $dX_t = f(X_t)\,dt$, $X_0 = x$, and extend to negative times t via $\xi^f_t(x) \stackrel{\text{def}}{=} \xi^{-f}_{-t}(x)$. Then

$$\frac{d\,\xi^f_t(x)}{dt} = f\big(\xi^f_t(x)\big) \quad \forall\, t \in (-\infty, +\infty) \;\; \text{, with } \; \xi^f_0(x) = x \; . \tag{5.1.23}$$

This is the ***flow generated*** by f on $\mathbb{R}^n$. Namely,

Proposition 5.1.8 *(i) For every $t \in \mathbb{R}$, $\xi^f_t : x \mapsto \xi^f_t(x)$ is a Lipschitz-continuous map from $\mathbb{R}^n$ to $\mathbb{R}^n$, and $t \mapsto \xi^f_t$ is a group under composition; i.e., for all $s, t \in \mathbb{R}$*

$$\xi^f_{t+s} = \xi^f_t \circ \xi^f_s \; .$$

(ii) In fact, every one of the maps $\xi^f_t : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable, and the $n\times n$-matrix $\big(D\xi^f_t[x]\big)^\mu_\nu \stackrel{\text{def}}{=} \partial\xi^{f\,\mu}_t(x)\big/\partial x^\nu$ of partial derivatives satisfies the following linear differential equation, obtained by formal differentiation of (5.1.23) in x:[6]

$$\frac{d\,D\xi^f_t[x]}{dt} = Df\big(\xi^f_t(x)\big) \cdot D\xi^f_t[x] \;,\; D\xi^f_0[x] = I_n \; . \tag{5.1.24}$$

Consider now two vector fields f and g. Their ***Lie bracket*** is the vector field[7]

$$[f,g](x) \stackrel{\text{def}}{=} Df(x)\cdot g(x) - Dg(x)\cdot f(x) \; ,$$

or[1]

$$[f,g]^\mu \stackrel{\text{def}}{=} f^\mu_{;\nu} g^\nu - g^\mu_{;\nu} f^\nu \; , \qquad \mu = 1, \ldots, n \; .$$

---

[6] I_n is the identity matrix on $\mathbb{R}^n$.

[7] Subscripts after semicolons denote partial derivatives, e.g., $f_{;\nu} \stackrel{\text{def}}{=} \frac{\partial f}{\partial x^\nu}$, $f_{;\mu\nu} \stackrel{\text{def}}{=} \frac{\partial^2 f}{\partial x^\nu \partial x^\mu}$. Einstein's convention is in force: summation over repeated indices in opposite positions is implied.

The fields f, g are said to ***commute*** if $[f,g]=0$. ***Their flows*** ξ^f, ξ^g are said to ***commute*** if

$$\xi^g_t \circ \xi^f_s = \xi^f_s \circ \xi^g_t \ , \qquad s,t \in \mathbb{R} \ .$$

Proposition 5.1.9 *The flows generated by f, g commute if and only if f, g do.*

Proof. We shall prove only the harder implication, the sufficiency, which is needed in theorem 5.4.23 on page 326. Assume then that $[f,g]=0$. The $\mathbb{R}^n$-valued path

$$\Delta_t \stackrel{\text{def}}{=} D\xi^f_t(x)\cdot g(x) - g\big(\xi^f_t(x)\big) \ , \qquad t \ge 0 \ ,$$

satisfies
$$\frac{d\,\Delta_t}{dt} = Df\big(\xi^f_t(x)\big)\cdot D\xi^f_t(x)\cdot g(x) - Dg\big(\xi^f_t(x)\big)\cdot f\big(\xi^f_t(x)\big)$$

as $[f,g]=0$:
$$= Df\big(\xi^f_t(x)\big)\cdot D\xi^f_t(x)\cdot g(x) - Df\big(\xi^f_t(x)\big)\cdot g\big(\xi^f_t(x)\big)$$
$$= Df\big(\xi^f_t(x)\big)\cdot \Delta_t \ .$$

Since $\Delta_0 = 0$, the unique global solution of this linear equation is $\Delta_{\boldsymbol{\cdot}} \equiv 0$,

whence $D\xi^f_t(x)\cdot g(x) = g\big(\xi^f_t(x)\big) \qquad \forall\, t \in \mathbb{R}$. (∗)

Fix a t and set $\Delta'_s \stackrel{\text{def}}{=} \xi^g_s\big(\xi^f_t(x)\big) - \xi^f_t\big(\xi^g_s(x)\big)$, $\qquad s \ge 0$.

Then
$$\frac{d\,\Delta'_s}{ds} = g\big(\xi^g_s\big(\xi^f_t(x)\big)\big) - D\xi^f_t\big(\xi^g_s(x)\big)\cdot g\big(\xi^g_s(x)\big)$$

by (∗):
$$= g\big(\xi^g_s\big(\xi^f_t(x)\big)\big) - g\big(\xi^f_t\big(\xi^g_s(x)\big)\big) \ ,$$

and so
$$|\Delta'_s| \le \int_0^s \Big| g\big(\xi^g_\sigma\big(\xi^f_t(x)\big)\big) - g\big(\xi^f_t\big(\xi^g_\sigma(x)\big)\big)\Big|\, d\sigma$$
$$\le L\cdot\int_0^s |\Delta'_\sigma|\, d\sigma \ .$$

By lemma A.2.35 $\Delta'_{\boldsymbol{\cdot}} \equiv 0$: the flows ξ^f and ξ^g commute. ∎

Now let $f_1, \ldots, f_d$ be vector fields on $\mathbb{R}^n$ that have bounded and Lipschitz partial derivatives and that commute with each other; and let $\xi^{f_1}, \ldots, \xi^{f_d}$ be their associated flows. For any $\boldsymbol{z} = (z^1, \ldots, z^d) \in \mathbb{R}^d$ let

$$\Xi^{\boldsymbol{f}}[\,\boldsymbol{\cdot}\,, \boldsymbol{z}] : \mathbb{R}^n \to \mathbb{R}^n$$

denote the composition in any order (see proposition 5.1.9) of $\xi^{f_1}_{z^1}, \ldots, \xi^{f_d}_{z^d}$.

Proposition 5.1.10 *(i)* $\Xi^{\boldsymbol{f}}$ *is a **differentiable action** of $\mathbb{R}^d$ on $\mathbb{R}^n$ in the sense that the maps $\Xi^{\boldsymbol{f}}[\,\boldsymbol{\cdot}\,, \boldsymbol{z}] : \mathbb{R}^n \to \mathbb{R}^n$ are differentiable and*

$$\Xi^{\boldsymbol{f}}[\,\boldsymbol{\cdot}\,, \boldsymbol{z}+\boldsymbol{z}'] = \Xi^{\boldsymbol{f}}[\,\boldsymbol{\cdot}\,, \boldsymbol{z}] \circ \Xi^{\boldsymbol{f}}[\,\boldsymbol{\cdot}\,, \boldsymbol{z}'] \ , \qquad \boldsymbol{z}, \boldsymbol{z}' \in \mathbb{R}^d \ .$$

$\Xi^{\boldsymbol{f}}$ *solves the initial value problem* $\Xi^{\boldsymbol{f}}[x,0] = x$,

$$\frac{\partial\, \Xi^{\boldsymbol{f}}[x,\boldsymbol{z}]}{\partial z^\theta} = f_\theta\big(\Xi^{\boldsymbol{f}}[x,\boldsymbol{z}]\big) \ , \qquad \theta = 1, \ldots, d \ .$$

(ii) For a given $\boldsymbol{z} \in \mathbb{R}^d$ *let* $\boldsymbol{z}_{\boldsymbol{.}} : [0, \tau] \to \mathbb{R}^d$ *be any continuous and piecewise continuously differentiable curve that connects the origin with* $\boldsymbol{z}$*:* $\boldsymbol{z}_0 = 0$ *and* $\boldsymbol{z}_\tau = \boldsymbol{z}$*. Then* $\Xi^{\boldsymbol{f}}[x, \boldsymbol{z}_{\boldsymbol{.}}]$ *is the unique (see item 5.1.2) solution of the initial value problem*[1]

$$x_{\boldsymbol{.}} = x + \int_0^{\boldsymbol{.}} f_\eta(x_\sigma) \frac{dz_\sigma^\eta}{d\sigma}\, d\sigma\,, \tag{5.1.25}$$

and consequently $\Xi^{\boldsymbol{f}}[x, \boldsymbol{z}]$ *equals the value* x_τ *at* τ *of that solution.*

In particular, for fixed $\boldsymbol{z} \in \mathbb{R}^d$ *set* $\tau \overset{\text{def}}{=} |\boldsymbol{z}|$*,* $\boldsymbol{z}_\sigma \overset{\text{def}}{=} \sigma \boldsymbol{z}/\tau$*, and* $f \overset{\text{def}}{=} f_\eta\, z^\eta/\tau$*. Then* $\Xi^{\boldsymbol{f}}[x, \boldsymbol{z}]$ *is the value* x_τ *at* τ *of the solution* $x_{\boldsymbol{.}}$ *of the ordinary initial value problem* $dx_\sigma = f(x_\sigma)\, d\sigma$*,* $x_0 = x$*:* $\Xi^{\boldsymbol{f}}[x, \boldsymbol{z}] = \xi[x, \tau; f]$*.*

ODE: Approximation

Picard's method constructs the solution $x_{\boldsymbol{.}}$ of equation (5.1.7),

$$x_t = c + \int_0^t f(x_s)\, ds\,, \tag{5.1.26}$$

as an iterated limit. Namely, every Picard iterate $x_{\boldsymbol{.}}^{(n)}$ is a limit, by virtue of being an integral, and $x_{\boldsymbol{.}}$ is the limit of the $x_{\boldsymbol{.}}^{(n)}$. As we have seen, this fact does not complicate questions of existence, uniqueness, and stability of the solution. It does, however, render nearly impossible the numerical computation of $x_{\boldsymbol{.}}$.

There is of course a plethora of approximation schemes that overcome this conundrum, from Euler's method of little straight steps to complex multistep methods of high accuracy. We give here a common description of most single-step methods. This is meant to lay down a foundation for the generalization in section 5.4 to the stochastic case, and to provide the classical results needed there. We assume for simplicity's sake that the initial condition is a constant $c \in \mathbb{R}^n$.

A ***single-step method*** is a procedure that from a ***threshold*** or ***step size*** δ and from the coefficient[8] f produces both a partition $0 = t_0 < t_1 < \ldots$ of time and a function

$$(x, t) \mapsto \xi'[x, t] = \xi'[x, t; f]$$

that has the following purpose: when the approximate solution x'_t has been constructed for $0 \le t \le t_k$, then ξ' is used to extend it to $[0, t_{k+1}]$ via

$$x'_t \overset{\text{def}}{=} \xi'[x'_{t_k}, t - t_k] \quad \text{for} \quad t_k \le t \le t_{k+1}\,. \tag{5.1.27}$$

(ξ' is typically evaluated only once per step, to compute the next point $x'_{t_{k+1}}$.) If the approximation scheme at hand satisfies this description, then we talk about ***the method*** $\boldsymbol{\xi'}$. If the t_k are set in advance, usually by $t_k \overset{\text{def}}{=} \delta \cdot k$,

[8] If the coupling coefficient depends explicitely on time, apply the time rectification of example 5.2.6.

then it is a ***non-adaptive method***; if the next time t_{k+1} is determined from δ, the situation at time t_k, and its outlook $\xi'[x_{t_k}, t-t_k]$ at that time, then the method ξ' is ***adaptive***. For instance, Euler's method of little straight steps is defined by $\xi'[x,t] = x + f(x)t$; it can be made adaptive by defining the stop for the next iteration as

$$t_{k+1} \stackrel{\text{def}}{=} \inf\{t > t_k : |\xi'[x'_{t_k}, t-t_k] - x'_{t_k}| \ge \delta\} \ :$$

"proceed to the next calculation only when the increment is large enough to warrant a new computation."

For the remainder of this short discussion of numerical approximation a non-adaptive single-step method ξ' is fixed. We shall say that ξ' has ***local order*** r on the coupling coefficient f if there exists a constant $\underline{m}$ such that[4,9] for $t \ge 0$

$$\big|\xi'[c,t;f] - \xi[c,t;f]\big| \le (|c|{+}1) \times (\underline{m}t)^r e^{\underline{m}t} \ . \tag{5.1.28}$$

The smallest such $\underline{m}$ will be denoted by $\underline{m}[f] = \underline{m}[f;\xi']$. If ξ' has local order r on all coefficients of class[10] C_b^∞, then it is simply said to have local order r. Inequality (5.1.28) will then usually hold on all coefficients of class C_b^k provided that k is sufficiently large.

We say that ξ' is of ***global order*** $r > 0$ on f if the difference of the exact solution $x_{\bullet} = \xi[c, \cdot\,; f]$ of (5.1.26) from its approximate $x'_{\bullet}$, made for the threshold δ via (5.1.27), satisfies an estimate $\|x'_{\bullet} - x_{\bullet}\|_{\overline{m}} = (|c|{+}1) \cdot O(\delta^r)$ for some constant $\overline{m} = \overline{m}[f;\xi']$. This amounts to saying that there exists a constant $\overline{b} = \overline{b}[f,\xi']$ such that

$$\big|x'_t - \xi[c,t;f]\big| \le \overline{b}\cdot(|c|{+}1) \times \delta^r e^{\overline{m}t}$$

for all sufficiently small $\delta > 0$, all $t \ge 0$, and all $c \in \mathbb{R}^n$. Euler's method for example is locally of order 2 on $f \in C_b^2$, and therefore is globally of order 1 according to the following criterion, whose proof is left to the reader:

Criterion 5.1.11 *(i) Suppose that the growth of ξ' is limited by the inequality $|\xi'[c,t]| \le C' \cdot (|c|{+}1)\, e^{M't}$, with C', M' constants. If $|\xi'[c,t;f] - \xi[c,t;f]| = (|c|{+}1) \cdot O(t^r)$ as $t \to 0$, then ξ' has local order r.* The usual Runge–Kutta and Taylor methods meet this description.

(ii) If the second-order mixed partials of ξ' are bounded on $\mathbb{R}^n \times [0,1]$, say by the constant $L' < \infty$, then, for $\delta \le 1$,

$$|\xi'[c',\cdot\,] - \xi'[c,\cdot\,]|_\delta^\star \le e^{L'\delta} \cdot |c'{-}c| \ .$$

(iii) If ξ' satisfies this inequality and has local order r, then it has global order $r{-}1$.

---

[9] This definition is not entirely standard – see however criterion 5.1.11 (i). In the present formulation, though, the notion is best used in, and generalized to, stochastic equations.

[10] A function f is of class C^k if it has continuous partial derivatives of order $1, \ldots, k$. It is of class C_b^k if it and these partials are bounded. One also writes $f \in C_b^k$; f is of class C_b^∞ if it is of class C_b^k for all $k \in \mathbb{N}$.

Note 5.1.12 ***Scaling*** provides a cheap way to produce new single-step methods from old. Here is how. With the substitutions $s \overset{\text{def}}{=} \alpha\sigma$ and $y_\sigma \overset{\text{def}}{=} x_{\alpha\sigma}$, equation (5.1.26) turns into

$$y_\tau = c + \int_0^\tau \alpha f(y_\sigma)\, d\sigma\ .$$

Now $\qquad |y_\tau - \xi'[c,\tau;\alpha f]| \le (|c|{+}1) \times (\underline{m}[\alpha f]\tau)^r\, e^{\underline{m}[\alpha f]\tau}$

begets $\qquad |x_t - \xi'[c,t/\alpha;\alpha f]| \le (|c|{+}1) \times \left(\frac{\underline{m}[\alpha f]}{\alpha}\cdot t\right)^r \times e^{\frac{\underline{m}[\alpha f]}{\alpha}\cdot t}\ .$

That is to say, $\xi'_\alpha : (c,t;f) \mapsto \xi'[c,t/\alpha;\alpha f]$ is another single-step method of local order $r+1$ and constant $\underline{m}[\alpha f]/\alpha$. If this constant is strictly smaller than $\underline{m}[f]$, then the new method ξ'_α is clearly preferable to ξ'.

It is easily seen that Taylor methods and Runge–Kutta methods are in fact ***scale-invariant*** in the sense that $\xi'_\alpha = \xi'$ for all $\alpha > 0$. The constant $\underline{m}[f;\xi']$ due to its minimality then equals the infimum over α of the constants $\underline{m}[\alpha f]/\alpha$, and this evidently has the following effect whenever the method ξ' has local order r on f:

If ξ' is scale-invariant, then $\underline{m}[\alpha f;\xi'] = \underline{m}[f;\xi']/\alpha$ for all $\alpha > 0$.

5.2 Existence and Uniqueness of the Solution

We shall in principle repeat the arguments of pages 274–281 to solve and estimate the general stochastic differential equation (5.1.2), which we recall here:[1]

$$X = C + F_\eta[X]_{.-}{*}Z^\eta \quad \text{, or} \quad X = C + \boldsymbol{F}[X]_{.-}{*}\boldsymbol{Z}\ . \tag{5.2.1}$$

To solve this equation we consider of course the map $\mathfrak{U}$ from the vector space $\mathfrak{D}^n$ of càdlàg adapted $\mathbb{R}^n$-valued processes to itself that is given by

$$\mathfrak{U}[X] \overset{\text{def}}{=} C + \boldsymbol{F}[X]_{.-}{*}\boldsymbol{Z}\ .$$

The problem (5.2.1) amounts to asking for a fixed point of $\mathfrak{U}$. As in the example of the previous section, its solution lies in designing complete norms[11] with respect to which $\mathfrak{U}$ is strictly contractive, Picard norms.

Henceforth the minimal assumptions (i)–(iii) of page 272 are in effect. In addition we will require throughout this and the next three sections that $\boldsymbol{Z} = (Z^1,\ldots,Z^d)$ is a local $L^q(\mathbb{P})$-integrator for some[12] $q \ge 2$ – except when this stipulation is explicitely rescinded on occasion.

[11] They are actually seminorms that vanish on evanescent processes, but we shall follow established custom and gloss over this point (see exercise A.2.31 on page 381).

[12] This requirement can of course always be satisfied provided we are willing to trade the given probability $\mathbb{P}$ for a suitable equivalent probability $\mathbb{P}'$ and to argue only up to some finite stopping time (see theorem 4.1.2). Estimates with respect to $\mathbb{P}'$ can be turned into estimates with respect to $\mathbb{P}$.

The Picard Norms

We will usually have selected a suitable exponent $p \in [2, q]$, and with it a norm $\| \ \|^*_{L^p}$ on random variables. To simplify the notation let us write

$$\left|\boldsymbol{F}^\nu\right|_\infty \stackrel{\text{def}}{=} \sup_{1\le\eta\le d} \left|F^\nu_\eta\right| \text{ , and } \left|\boldsymbol{F}\right|_{\infty p} \stackrel{\text{def}}{=} \Bigl(\sum_{1\le\nu\le n} \left|\boldsymbol{F}^\nu\right|^p_\infty\Bigr)^{1/p} \tag{5.2.2}$$

for the size of a d-tuple $\boldsymbol{F} = (F_1, \ldots, F_d)$ of n-vectors. Recall also that the maximal process of a vector $X \in \mathfrak{D}^n$ is the vector composed of the maximal functions of its components[13].

In the ordinary differential equation of page 274 both driver and controller were the same, to wit, time. In the presence of several drivers a common controller should be found and used to clock a common time transformation. THE strictly increasing previsible controller $\Lambda = \Lambda^{\langle q\rangle}[\boldsymbol{Z}]$ of theorem 4.5.1 and THE associated continuous time transformation $T^\bullet$ by predictable stopping times

$$T^\lambda \stackrel{\text{def}}{=} \inf\{t : \Lambda_t \ge \lambda\} \tag{5.2.3}$$

of remark 4.5.2 come to mind.[14] Since $\Lambda_t \ge \alpha \cdot t$, the T^λ are bounded, so that a negligible set in $\mathcal{F}_{T^\lambda}$ is nearly empty. Since $\Lambda_t < \infty \quad \forall\, t$, the T^λ increase without bound as $\lambda \to \infty$.

We use THE time transformation and (5.2.2) to define, for any $p \in [2, q]$ and any $M \ge 1$, functionals $\|\ \|_{p,M}$ and $\|\ \|^\star_{p,M}$, the ***Picard norms***[11]

on vectors[2] $\qquad X = (X^1, \ldots, X^n) \in \mathfrak{D}^n$

by $$\|X\|_{p,M} \stackrel{\text{def}}{=} \sup_{\lambda>0} e^{-M\lambda} \cdot \Bigl\| \left|X_{T^\lambda-}\right|_p \Bigr\|^*_{L^p(\mathbb{P})} ,$$

which is less than $$\|X\|^\star_{p,M} \stackrel{\text{def}}{=} \sup_{\lambda>0} e^{-M\lambda} \cdot \Bigl\| \left|X^\star_{T^\lambda-}\right|_p \Bigr\|^*_{L^p(\mathbb{P})} . \tag{5.2.4}$$

Then we set $$\mathfrak{S}^n_{p,M} \stackrel{\text{def}}{=} \Bigl\{ X \in \mathfrak{D}^n \ :\ \|X\|_{p,M} < \infty \Bigr\} ,$$

which clearly contains $$\mathfrak{S}^{\star n}_{p,M} \stackrel{\text{def}}{=} \Bigl\{ X \in \mathfrak{D}^n \ :\ \|X\|^\star_{p,M} < \infty \Bigr\} .$$

For the meaning of $\|\ \|^*_{L^p(\mathbb{P})}$ see item A.8.21 on page 452; it is used instead of $\|\ \|_{L^p(\mathbb{P})}$ to avoid worries about finiteness and measurability of its argument. Definition (5.2.4) is a straighforward generalization of (5.1.10).

The next lemma illuminates the role of the functionals (5.2.4) in the control of the driver $\boldsymbol{Z}$:

---

[13] $X^\star \stackrel{\text{def}}{=} \bigl(|X^1|^\star, \ldots |X^n|^\star\bigr)$ has size $|X|^\star_p \le |X^\star|_p \le n^{1/p} \cdot |X|^\star_p$.

[14] This is disingenuous; Λ and $T^\bullet$ were of course specifically designed for the problem at hand.

Lemma 5.2.1 *(i)* $\lVert\ \rVert_{p,M}$ *and* $\lVert\ \rVert^\star_{p,M}$ *are seminorms on* $\mathfrak{S}^n_{p,M}$ *and* $\mathfrak{S}^{\star n}_{p,M}$*, respectively.* $\mathfrak{S}^{\star n}_{p,M}$ *is complete under* $\lVert\ \rVert^\star_{p,M}$*. A process* X *has Picard norm*[11] $\lVert X\rVert^\star_{p,M} = 0$ *if and only if it is evanescent.*

(ii) $\lVert X\rVert^\star_{p,M}$ *increases as* p *increases and as* M *decreases and*

$$\lVert X\rVert^\star_{p,M} = \sup\left\{\lVert X\rVert^\star_{p^\circ,M^\circ} : p^\circ < p, M^\circ > M\right\}, \qquad X \in \mathfrak{D}^n .$$

(iii) On any d-tuple $\boldsymbol{F} = (F_1, \ldots, F_d)$ *of adapted càdlàg* $\mathbb{R}^n$*-valued processes*

$$\lVert \boldsymbol{F}_{\bullet-} {*} \boldsymbol{Z}\rVert^\star_{p,M} \le \frac{C_p^{\diamond(4.5.1)}}{M^{1/p^\diamond}} \cdot \lVert |\boldsymbol{F}|_\infty \rVert_{p,M} . \tag{5.2.5}$$

Proof. Parts (i) and (ii) are quite straightforward and are left to the reader.

(iii) Pick a $\mu > 0$ and let S be a stopping time that is strictly prior to T^μ on $[T^\mu > 0]$. Fix a $\nu \in \{1, \ldots, n\}$. With that, theorem 4.5.1 gives

$$\left\lVert \left(\boldsymbol{F}^\nu_{\bullet-} {*} \boldsymbol{Z}\right)^\star_S \right\rVert^*_{L^p} \le C_p^{\diamond(4.5.1)} \cdot \max_{\rho=1^\diamond, p^\diamond} \left\lVert \left(\int_0^S \left|\boldsymbol{F}^\nu_{s-}\right|^\rho_\infty d\Lambda_s\right)^{1/\rho} \right\rVert^*_{L^p(\mathbb{P})}$$

by theorem 2.4.7:
$$\le C_p^\diamond \cdot \max_\rho \left\lVert \left(\int_{[T^\lambda \le S]} \left|\boldsymbol{F}^\nu_{T^\lambda -}\right|^\rho_\infty d\lambda\right)^{1/\rho} \right\rVert^*_{L^p(\mathbb{P})}$$

(!)
$$\le C_p^\diamond \cdot \max_\rho \left\lVert \left(\int_{[\lambda \le \mu]} \left|\boldsymbol{F}^\nu_{T^\lambda -}\right|^\rho_\infty d\lambda\right)^{1/\rho} \right\rVert^*_{L^p(\mathbb{P})} . \tag{!}$$

The previsibility of the controller Λ is used in an essential way at (!). Namely, $T^\lambda \le T^\mu$ does not in general imply $\lambda \le \mu$ – in fact, λ could exceed μ by as much as the jump of Λ at T^μ. However, $T^\lambda < T^\mu$ does imply $\lambda < \mu$, and we produce this inequality by calculating only up to a stopping time S *strictly prior* to T^μ. That such exist arbitrarily close to T^μ is due to the previsibility of Λ, which implies the predictability of T^μ (exercise 3.5.19). We use this now: letting the S above run through a sequence announcing T^μ yields

$$\left\lVert \left(\boldsymbol{F}^\nu_{\bullet-} {*} \boldsymbol{Z}\right)^\star_{T^\mu -} \right\rVert^*_{L^p(\mathbb{P})} \le C_p^\diamond \cdot \max_{\rho=1^\diamond, p^\diamond} \left\lVert \left(\int_0^\mu \left|\boldsymbol{F}^\nu_{T^\lambda -}\right|^\rho_\infty d\lambda\right)^{1/\rho} \right\rVert^*_{L^p(\mathbb{P})} . \tag{5.2.6}$$

Applying the ℓ^p-norm $|\ |_p$ to these n-vectors and using Fubini produces

$$\left\lVert \left|\left(\boldsymbol{F}_{\bullet-} {*} \boldsymbol{Z}\right)^\star_{T^\mu -}\right|_p \right\rVert^*_{L^p(\mathbb{P})} \le C_p^\diamond \cdot \max_\rho \left\lVert \left|\left(\int_0^\mu \left|\boldsymbol{F}_{T^\lambda -}\right|^\rho_\infty d\lambda\right)^{1/\rho}\right|_p \right\rVert^*_{L^p(\mathbb{P})}$$

by exercise A.3.29:
$$\le C_p^\diamond \cdot \max_\rho \left(\int_0^\mu \left\lVert \left|\boldsymbol{F}_{T^\lambda -}\right|_{\infty p} \right\rVert^{*\rho}_{L^p(\mathbb{P})} d\lambda\right)^{1/\rho}$$

by definition (5.2.4):
$$\le C_p^\diamond \cdot \max_\rho \left(\int_0^\mu \lVert |\boldsymbol{F}|_\infty \rVert^\rho_{p,M} \cdot e^{M\lambda\rho}\, d\lambda\right)^{1/\rho}$$

$$= C_p^\diamond \cdot \lVert |\boldsymbol{F}|_\infty \rVert_{p,M} \cdot \max_\rho \left(\int_0^\mu e^{M\lambda\rho}\, d\lambda\right)^{1/\rho}$$

with a little calculus: $$< C_p^\diamond \cdot \big\| |\boldsymbol{F}|_\infty \big\|_{p,M} \cdot \max_{\rho=1^\diamond,p^\diamond} \frac{e^{M\mu}}{(M\rho)^{1/\rho}}$$

since $M \geq 1 \leq \rho^{1/\rho}$: $$\leq \frac{C_p^\diamond\, e^{M\mu}}{M^{1/p^\diamond}} \cdot \big\| |\boldsymbol{F}|_\infty \big\|_{p,M} .$$

Multiplying this by $e^{-M\mu}$ and taking the supremum over $\mu > 0$ results in inequality (5.2.5).

Note here that the use of a sequence announcing T^μ provides information only about the left-continuous version of $\big(\boldsymbol{F}_{.-}*\boldsymbol{Z}\big)^\star$ at T^μ; this explains why we chose to define $\|X\|_{p,M}$ and $\|X\|^\star_{p,M}$ using the left-continuous versions $X_{.-}$ and $X^\star_{.-}$ rather than $X_.$ and $X^\star_.$ itself. However, if $\boldsymbol{Z}$ is quasi-left-continuous and therefore Λ is continuous (exercise 4.5.16) and $T^\cdot$ strictly increasing,[15] then we can define $\|\ \|_{p,M}$ and $\|\ \|^\star_{p,M}$ using $X^\star_.$ itself, with inequality (5.2.5) persisting. In fact, the computation leading to this inequality then simplifies a little, since we can take $S = T^\mu$ to start with. ∎

Here are further useful facts about the functionals $\|\ \|_{p,M}$ and $\|\ \|^\star_{p,M}$:

Exercise 5.2.2 Let $\Delta_\eta \in \mathfrak{L}$ with $|\Delta_\eta| \leq \delta \in \mathbb{R}_+$. Then $\|\Delta_\eta * Z^\eta\|^\star_{p,M} \leq \delta C_p^\diamond / eM$.

Exercise 5.2.3 Let $p \in [2, q]$ and $M > 0$. The seminorm

$$X \mapsto \|X\|^\bullet_{p,M} \overset{\text{def}}{=} \Big(\sum_{1\leq\nu\leq n} \int_0^\infty \|X^{\nu\star}_{T^\lambda-}\|^{*p}_{L^p}\ Me^{-M\lambda} d\lambda \Big)^{1/p}$$

is complete on $$\mathfrak{S}^\bullet_{p,M} \overset{\text{def}}{=} \Big\{ X \in \mathfrak{D}^n : \|X\|^\bullet_{p,M} < \infty \Big\}$$

and satisfies the following analog of inequality (5.2.5):

$$\|\boldsymbol{F}_{.-} * \boldsymbol{Z}\|^\bullet_{p,M} \leq C_p^{\diamond(4.5.1)} \Big(\frac{1}{M^{1/p}} \vee \frac{p}{M} \Big) \cdot \big\| |\boldsymbol{F}|_\infty \big\|^\bullet_{p,M} .$$

Furthermore, $p \mapsto \|X\|^\bullet_{p,M}$ is increasing, $M \mapsto \|X\|^\bullet_{p,M}$ is decreasing, and

$$\|X\|^\bullet_{p,M'} \leq \Big(\frac{M'}{M' - Mp} \Big)^{1/p} \|X\|^\star_{p,M} \qquad \text{for } M' > Mp ,$$

and $$\|X\|^\star_{p,M} \leq \|X\|^\bullet_{p,M'} \qquad \text{for } M' \leq Mp .$$

The seminorms $\|\ \|^\bullet_{p,M}$ are just as good as the $\|\ \|^\star_{p,M}$ for the development.

Lipschitz Conditions

As above in section 5.1 on ODEs, Lipschitz conditions on the coupling coefficient $\boldsymbol{F}$ are needed to solve (5.2.1) with Picard's scheme. A rather restrictive

[15] This happens for instance when $\boldsymbol{Z}$ is a Lévy process or the solution of a stochastic differential equation driven by a Lévy process (see exercise 5.2.17).

one is this ***strong Lipschitz condition***: there exists a constant $L < \infty$ such that for any two $X, Y \in \mathfrak{D}^n$

$$\left|\left(\boldsymbol{F}[Y] - \boldsymbol{F}[X]\right)\right|_{\infty p} \le L \cdot \left|Y - X\right|_p \tag{5.2.7}$$

up to evanescence. It clearly implies the slightly weaker Lipschitz condition

$$\left|\left(\boldsymbol{F}[Y] - \boldsymbol{F}[X]\right)_{\cdot-}\right|_{\infty p} \le L \cdot \left|\left(Y - X\right)^\star\right|_p , \tag{5.2.8}$$

which is to say that at any finite stopping time T almost surely

$$\Big(\sum_\nu \sup_\eta \left|\left(F^\nu_\eta[Y] - F^\nu_\eta[X]\right)_{T-}\right|^p\Big)^{1/p} \le L \cdot \Big(\sum_\nu \sup_{s \le T} |Y^\nu - X^\nu|^p_s\Big)^{1/p} .$$

These conditions are independent of p in the sense that if one of them holds for one exponent $p \in (0, \infty]$, then it holds for any other, except that the Lipschitz constant may change with p. Inequality (5.2.8) implies the following rather much weaker ***p-mean Lipschitz condition:***

$$\Big\|\left|\left(\boldsymbol{F}[Y] - \boldsymbol{F}[X]\right)_{T-}\right|_{\infty p}\Big\|_{L^p(\mathbb{P})} \le L \cdot \Big\|\left|\left(Y - X\right)^\star_t\right|_p\Big\|_{L^p(\mathbb{P})} , \tag{5.2.9}$$

which in turn implies that at any predictable stopping time T

$$\Big\|\left|\left(\boldsymbol{F}[Y] - \boldsymbol{F}[X]\right)_{T-}\right|_{\infty p}\Big\|_{L^p(\mathbb{P})} \le L \cdot \Big\|\left|\left(Y - X\right)^\star_{T-}\right|_p\Big\|_{L^p(\mathbb{P})} , \tag{5.2.10}$$

in particular for the stopping times T^λ of THE time transformation (5.2.3)

$$\Big\|\left|\left(\boldsymbol{F}[Y] - \boldsymbol{F}[X]\right)_{T^\lambda-}\right|_{\infty p}\Big\|_{L^p(\mathbb{P})} \le L \cdot \Big\|\left|\left(Y - X\right)^\star_{T^\lambda-}\right|_p\Big\|_{L^p(\mathbb{P})} , \tag{5.2.11}$$

whenever $X, Y \in \mathfrak{D}^n$. Inequality (5.2.10) can be had simply by applying (5.2.9) to a sequence that announces T and taking the limit. Finally, multiplying (5.2.11) with $e^{-M\lambda}$ and taking the supremum over $\lambda > 0$ results in

$$\left\| \left|\boldsymbol{F}[Y] - \boldsymbol{F}[X]\right|_\infty \right\|_{p,M} \le L \cdot \left\| X - Y \right\|^\star_{p,M} \tag{5.2.12}$$

for $X, Y \in \mathfrak{D}^n$. This is the form in which any Lipschitz condition will enter the existence, uniqueness, and stability proofs below. If it is satisfied, we say that $\boldsymbol{F}$ is ***Lipschitz in the Picard norm.***[11] See remark 5.2.20 on page 294 for an example of a coupling coefficient that is Lipschitz in the Picard norm without being Lipschitz in the sense of inequality (5.2.8).

The adjusted coupling coefficient ${}^0\boldsymbol{F}$ of equation (5.1.6) shares any or all of these Lipschitz conditions with $\boldsymbol{F}$, and any of them guarantees the non-anticipating nature (5.1.4) of $\boldsymbol{F}$ and of ${}^0\boldsymbol{F}$, at least at the stopping times entering their definition.

Here are a few examples of coupling coefficients that are strongly Lipschitz in the sense of inequality (5.2.7) and therefore also in the weaker senses of (5.2.8)–(5.2.12). The verifications are quite straightforward and are left to the reader.

Example 5.2.4 Suppose the F_η are ***markovian***, that is to say, are of the form $F_\eta[X] = f_\eta \circ X$, with $f_\eta : \mathbb{R}^n \to \mathbb{R}^n$ vector fields. If the f_η are ***Lipschitz***, meaning that[4]

$$\big| f_\eta(x) - f_\eta(y) \big| \le L_\eta \cdot |x - y| \tag{5.2.13}$$

for some constants L_η and all $x, y \in \mathbb{R}^n$, then $\boldsymbol{F}$ is Lipschitz in the sense of (5.2.7). This will be the case in particular when the partial derivatives[16] $f^\mu_{\eta;\nu}$ exist and are bounded for every η, ν, μ. Most Lipschitz coupling coefficients appearing at present in physical models, financial models, etc., are of this description. They are used when only the current state X_t of X influences its evolution, when information of how it got there is irrelevant. Markovian coefficients are also called ***autonomous***.

Example 5.2.5 In a slight generalization, we call F_η an ***instantaneous coupling coefficient*** if there exists a Borel vector field $f_\eta : [0,\infty)\times\mathbb{R}^n \to \mathbb{R}^n$ so that $F_\eta[X]_s = f_\eta(s, X_s)$ for $s \in [0,\infty)$ and $X \in \mathfrak{D}^n$. If f_η is ***equi-Lipschitz in its spacial arguments***, meaning that[4]

$$\sup_s \big| f_\eta(s,x) - f_\eta(s,y) \big| \le L_\eta \cdot |x - y|$$

for some constants L_η and all $x, y \in \mathbb{R}^n$, then $\boldsymbol{F}$ is strongly Lipschitz in the sense of (5.2.7). A markovian coupling coefficient clearly is instantaneous.

Example 5.2.6 (Time Rectification of Instantaneous Equations) The two previous examples are actually not too far apart. Suppose the instantaneous coefficients $(s,x) \mapsto f_\eta(s,x)$ happen to be Lipschitz in all of their arguments, which is to say

$$\big| f_\eta(s,x) - f_\eta(t,y) \big| \le L_\eta \cdot \big|(s,x) - (t,y)\big| \,. \tag{5.2.14}$$

Then we expand the driver by giving it the zeroth component $Z^0_t = t$ and set $\quad \boldsymbol{Z}^\vdash \overset{\text{def}}{=} (Z^0, Z^1, \ldots, Z^d)$,

expand the state space from $\mathbb{R}^n$ to $\mathbb{R}^{n+1} = (-\infty,\infty) \times \mathbb{R}^n$,

setting $\quad X^\vdash \overset{\text{def}}{=} (X^0, X^1, \ldots, X^n)$,

expand the initial state to

$$C^\vdash \overset{\text{def}}{=} (0, C^1, \ldots, C^n) \,,$$

and consider the expanded *and now markovian* differential equation

$$\begin{pmatrix} X^0 \\ X^1 \\ \vdots \\ X^n \end{pmatrix} = \begin{pmatrix} C^0 \\ C^1 \\ \vdots \\ C^n \end{pmatrix} + \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & f^1_1(X^\vdash)_{.-} & \cdots & f^1_d(X^\vdash)_{.-} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & f^n_1(X^\vdash)_{.-} & \cdots & f^n_d(X^\vdash)_{.-} \end{pmatrix} * \begin{pmatrix} Z^0 \\ Z^1 \\ \vdots \\ Z^d \end{pmatrix}$$

$$\text{or} \qquad X^\vdash = C^\vdash + \boldsymbol{f}^\vdash(X^\vdash)_{.-} * \boldsymbol{Z}^\vdash \,,$$

[16] Subscripts after semicolons denote partial derivatives, e.g., $f_{\eta;\nu} \overset{\text{def}}{=} \frac{\partial f_\eta}{\partial x^\nu}$, $f_{\eta;\mu\nu} \overset{\text{def}}{=} \frac{\partial^2 f_\eta}{\partial x^\nu \partial x^\mu}$.

in obvious notation. The first line of this equation simply reads $X^0_t = t$; the others combine to the original equation $X_t = C_t + \sum_{\eta=1}^d \int_0^t f_\eta(s, X_s)\, dZ^\eta_s$. In this way it is possible to generalize very cheaply results about markovian stochastic differential equations to instantaneous stochastic differential equations, at least in the presence of inequality (5.2.14).

Example 5.2.7 We call F an ***autologous coupling coefficient*** if there exists an adapted map[17] $f : \mathscr{D}^n \to \mathscr{D}^n$ so that $F[X]_{\bullet}(\omega) = f\big[X_{\bullet}(\omega)\big]$ for nearly all $\omega \in \Omega$. We say that such f is ***Lipschitz*** with constant L if[4] for any two paths $x_{\bullet}, y_{\bullet} \in \mathscr{D}^n$ and all $t \geq 0$

$$\big|\, f\big[x_{\bullet}\big] - f\big[y_{\bullet}\big]\,\big|_{t-} \leq L \cdot |x_{\bullet} - y_{\bullet}|^{\star}_t \,. \tag{5.2.15}$$

In this case the coupling coefficient F is evidently strongly Lipschitz, and thus is Lipschitz in any of the Picard norms as well. Autologous[18] coupling coefficients might be used to model the influence of the whole past of a path of $X_{\bullet}$ on its future evolution. Instantaneous autologous coupling coefficients are evidently autonomous.

Example 5.2.8 A particular instance of an autologous Lipschitz coupling coefficient is this: let $v = (v^\nu_\mu)$ be an $n{\times}n$-matrix of deterministic càdlàg functions on the half-line that have uniformly bounded total variation, and let v act by convolution:

$$F^\nu[X]_t(\omega) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \textstyle\sum_\mu X^\mu_{t-s}(\omega)\; dv^\nu_{\mu s} = \displaystyle\int_0^t \textstyle\sum_\mu X^\mu_{t-s}(\omega)\; dv^\nu_{\mu s} \,.$$

(As usual we think $X_s = v_s = 0$ for $s < 0$.) Such autologous Lipschitz coupling coefficients could model systematic influences of the past of X on its evolution that abate as time elapses. Technical stock analysts who believe in trends might use such coupling coefficients to model the evolution of stock prices.

Example 5.2.9 We call F a ***randomly autologous coupling coefficient*** if there exists a function $f : \Omega \times \mathscr{D}^n \to \mathscr{D}^n$, adapted to $\mathcal{F}_{\bullet} \otimes \mathcal{F}_{\bullet}[\mathscr{D}^n]$, such that $F[X]_{\bullet}(\omega) = f\big[\omega, X_{\bullet}(\omega)\big]$ for nearly all $\omega \in \Omega$. We say that such f is ***Lipschitz*** with constant L if[4] for any two paths $x_{\bullet}, y_{\bullet} \in \mathscr{D}^n$ and all $t \geq 0$

$$\big|\, f\big[\omega, x_{\bullet}\big] - f\big[\omega, y_{\bullet}\big]\,\big|_{t-} \leq L \cdot |x_{\bullet} - y_{\bullet}|^{\star}_t \tag{5.2.16}$$

at nearly every $\omega \in \Omega$. In this case the coupling coefficient F is evidently strongly Lipschitz, and thus is Lipschitz in any of the Picard norms as well. Here are several examples of randomly autologous coupling coefficients:

[17] See item 2.3.8. We may equip path space with its canonical or its natural filtration, *ad libitum*; consistency behooves, however.
[18] To the choice of word: if at the time of the operation the patient's own blood is used, usually collected previously on many occasions, then one talks of an autologous blood supply.

Example 5.2.10 Let $D = (D^\nu_\mu)$ be an $n \times n$-matrix of uniformly bounded adapted càdlàg processes. Then $X \mapsto \sum_\mu D^\nu_\mu X^\mu$ is Lipschitz in the sense of (5.2.7). Such coupling coefficients appear automatically in the stability theory of stochastic differential equations, even of those that start out with markovian coefficients (see section 5.3 on page 298). They would generally be used to model random influences that the past of X has on its future evolution.

Example 5.2.11 Let $V = (V^\nu_\mu)$ be a matrix of adapted càdlàg processes that have bounded total variation. Define F by

$$F^\nu[X]_t(\omega) \stackrel{\text{def}}{=} \int_0^t \sum_\mu X^\mu_s(\omega)\, dV^\nu_{\mu s}(\omega) \; .$$

Such F is evidently randomly autologous and might again be used to model random influences that the past of X has on its future evolution.

Example 5.2.12 We call F an ***endogenous coupling coefficient*** if there exists an adapted function[17] $f : \mathscr{D}^d \times \mathscr{D}^n \to \mathscr{D}^n$ so that

$$F[X]_{\bullet}(\omega) = f\big[\boldsymbol{Z}_{\bullet}(\omega), X_{\bullet}(\omega)\big]$$

for nearly all $\omega \in \Omega$. We say that such f is ***Lipschitz*** with constant L if[4] for any two paths $x_{\bullet}, y_{\bullet} \in \mathscr{D}^n$ and all $\boldsymbol{z}_{\bullet} \in \mathscr{D}^d$ and $t \geq 0$

$$\big| f\big[\boldsymbol{z}_{\bullet}, x_{\bullet}\big] - f\big[\boldsymbol{z}_{\bullet}, y_{\bullet}\big]\big|_{t-} \leq L \cdot |x_{\bullet} - y_{\bullet}|^\star_t \; . \tag{5.2.17}$$

In this case the coupling coefficient F is evidently strongly Lipschitz, and thus is Lipschitz also in any of the Picard norms. Autologous coupling coefficients are evidently endogenous. Conversely, simply adding the equation $Z^\eta = \delta^\eta_\theta * Z^\theta$ to (5.2.1) turns that equation into an autologous equation for the vector $(X, \boldsymbol{Z}) \in \mathfrak{D}^{n+d}$. Equations with endogenous coefficients can be solved numerically by an algorithm (see theorem 5.4.5 on page 316).

Example 5.2.13 (Permanence Properties) If F, F' are strongly Lipschitz coupling coefficients with $d = 1$, then so is their composition. If the $F_1, F_2, \ldots$ each are strongly Lipschitz with $d = 1$, then the finite collection $\boldsymbol{F} \stackrel{\text{def}}{=} (F_1, \ldots, F_d)$ is strongly Lipschitz.

Existence and Uniqueness of the Solution

Let us now observe how our Picard norms[11] (5.2.4) and the Lipschitz condition (5.2.12) cooperate to produce a solution of equation (5.2.1), which we recall as

$$X = C + F_\eta[X]_{\bullet-} * Z^\eta \; ; \tag{5.2.18}$$

i.e., how they furnish a fixed point of

$$\mathfrak{U} : X \mapsto C + \boldsymbol{F}[X]_{\bullet-} * \boldsymbol{Z} \; .$$

We are of course after the contractivity of $\mathfrak{U}$.

To establish it consider two elements X, Y in $\mathfrak{S}^{\star n}_{p,M}$ and estimate:

$$\left\|\mathfrak{U}[Y]-\mathfrak{U}[X]\right\|^{\star}_{p,M} = \left\|(\boldsymbol{F}[Y]-\boldsymbol{F}[X])_{.-}{*}\boldsymbol{Z}\right\|^{\star}_{p,M}$$

by inequality (5.2.5):
$$\leq \frac{C^{\diamond}_{p}}{M^{1/p^{\diamond}}}\cdot\left\|\left|\boldsymbol{F}[Y]-\boldsymbol{F}[X]\right|_{\infty}\right\|_{p,M}$$

by inequality (5.2.12):
$$\leq \frac{LC^{\diamond}_{p}}{M^{1/p^{\diamond}}}\cdot\left\|Y-X\right\|^{\star}_{p,M} . \tag{5.2.19}$$

Thus $\mathfrak{U}$ is strictly contractive provided M is sufficiently large, say

$$M > M^{\diamond}_{p,L} \stackrel{\text{def}}{=} \left(C^{\diamond}_{p}\, L\right)^{p^{\diamond}} . \tag{5.2.20}$$

$\mathfrak{U}$ then has modulus of contractivity

$$\gamma = \gamma_{p,M,L} \stackrel{\text{def}}{=} \left(M^{\diamond}_{p,L}/M\right)^{1/p^{\diamond}} \tag{5.2.21}$$

strictly less than 1. The arguments of items 5.1.3–5.1.5, adapted to the present situation, show that $\mathfrak{S}^{\star n}_{p,M}$ contains a unique fixed point X of $\mathfrak{U}$

provided
$$^{0}C \stackrel{\text{def}}{=} C+\boldsymbol{F}[0]_{.-}{*}\boldsymbol{Z} \in \mathfrak{S}^{\star n}_{p,M} , \tag{5.2.22}$$

which is the same as saying that $\quad \mathfrak{U}[0] \in \mathfrak{S}^{\star n}_{p,M}$;

and then they even furnish a priori estimates of the size of the solution X and its deviation from the initial condition, namely,

$$\left\|X\right\|^{\star}_{p,M} \leq \frac{1}{1-\gamma}\cdot\left\|{}^{0}C\right\|^{\star}_{p,M} , \tag{5.2.23}$$

and
$$\left\|X-C\right\|^{\star}_{p,M} \leq \frac{1}{1-\gamma}\cdot\left\|\boldsymbol{F}[C]_{.-}{*}\boldsymbol{Z}\right\|^{\star}_{p,M} \tag{5.2.24}$$

by inequality (5.2.5):
$$\leq \frac{C^{\diamond}_{p}}{(1-\gamma)M^{1/p^{\diamond}}}\cdot\left\|\boldsymbol{F}[C]\right\|^{\star}_{p,M} . \tag{5.2.25}$$

Alternatively, by solving equation (5.2.21) for M we may specify a modulus of contractivity $\gamma \in (0,1)$ in advance:

if we set
$$M_{L:\gamma} \stackrel{\text{def}}{=} (10qL/\gamma)^{q} , \tag{5.2.26}$$

then
$$M_{L:\gamma} \geq M^{\diamond(5.2.20)}_{p,L}/\gamma^{p^{\diamond}} ,$$

and (5.2.5) turns into
$$\left\|\boldsymbol{F}_{.-}{*}\boldsymbol{Z}\right\|^{\star}_{p,M} \leq \frac{\gamma}{L}\cdot\left\||\boldsymbol{F}|_{\infty}\right\|_{p,M} \tag{5.2.27}$$
$$\leq \frac{\gamma}{L}\cdot\left\||\boldsymbol{F}|_{\infty}\right\|^{\star}_{p,M} ,$$

and (5.2.19) into
$$\left\|\mathfrak{U}[Y]-\mathfrak{U}[X]\right\|^{\star}_{p,M} \leq \gamma\cdot\left\|Y-X\right\|^{\star}_{p,M}$$

for all $p \in [2,q]$ and all $M \geq M_{L:\gamma}$ simultaneously. To summarize:

Proposition 5.2.14 *In addition to the minimal assumptions **(ii)–(iii)** on page 272 assume that* $\boldsymbol{Z}$ *is a local* L^q*-integrator for some* $q \geq 2$ *and that* $\boldsymbol{F}$ *satisfies*[19] *the Picard norm Lipschitz condition (5.2.12) for some* $p \in [2, q]$ *and some* $M > M_{p,L}^{\diamond(5.2.20)}$. *If*

$$ {}^0C \stackrel{\mathrm{def}}{=} C + \boldsymbol{F}[0]_{.-}{*}\boldsymbol{Z} \quad \textit{belongs to} \quad \mathfrak{S}_{p,M}^{\star n} , \tag{5.2.28} $$

then $\mathfrak{S}_{p,M}^{\star n}$ *contains one and only one strong global solution* X *of the stochastic differential equation (5.2.18).*

We are now in position to establish a rather general existence and uniqueness theorem for stochastic differential equations, without assuming more about $\boldsymbol{Z}$ than that it be an L^0-integrator:

Theorem 5.2.15 *Under the minimal assumptions (i)–(iii) on page 272 and the strong Lipschitz condition (5.2.8) there exists a strong global solution* X *of equation (5.2.18), and up to indistinguishability only one.*

Proof. Note first that (5.2.8) for some $p > 0$ implies (5.2.8) and then (5.2.10) for any probability equivalent with $\mathbb{P}$ and for any $p \in (0, \infty)$, in particular for $p = 2$, except that the Lipschitz L constant may change as p is altered. Let U be a finite stopping time. There is a probability $\mathbb{P}'$ equivalent to the given probability $\mathbb{P}$ such that the 2+d stopped processes $|C^{\star U}|_2$, $\left|\left(\boldsymbol{F}[0]_{.-}{*}\boldsymbol{Z}\right)^{\star U}\right|_2$, and $Z^{\eta U}$, $\eta = 1 \ldots d$, are global $L^2(\mathbb{P}')$-integrators. Namely, all of these processes are global L^0-integrators (proposition 2.4.1 and proposition 2.1.9), and theorem 4.1.2 provides $\mathbb{P}'$. According to proposition 3.6.20, if X satisfies

$$ X = C^U + \boldsymbol{F}[X]_{.-}{*}\boldsymbol{Z}^U \tag{5.2.29} $$

in the sense of the stochastic integral with respect to $\mathbb{P}$, it satisfies the same equation in the sense of the stochastic integral with respect to $\mathbb{P}'$, and vice versa: as long as we want to solve the stopped equation (5.2.29) we might as well assume that $|C^{\star U}|_2$, $\left|\left(\boldsymbol{F}[0]_{.-}{*}\boldsymbol{Z}\right)^{\star U}\right|_2$, and $\boldsymbol{Z}^U$ are global L^2-integrators. Then condition (5.2.22) is clearly satisfied, whatever $M > 0$ (lemma 5.2.1 (iii)). We apply inequality (5.2.19) with $p = q = 2$ and $M > M_{2,L}^{\diamond(5.2.20)}$ to make $\mathfrak{U}$ strictly contractive and see that there is a solution of (5.2.29). Suppose there are two solutions X, X'. Then we can choose $\mathbb{P}'$ so that in addition the difference $X - X'$, which stops after time U, is a global $L^2(\mathbb{P}')$-integrator as well and thus belongs to $\mathfrak{S}_{2,M}^{\star n}(\mathbb{P}')$. Since the strictly contractive map $\mathfrak{U}$ has at most one fixed point, we must have $\| X - X' \|_{2,M}^{\star} = 0$, which means that X and X' are indistinguishable. Let X^U denote the unique solution of equation (5.2.29). We let U run through a sequence (U_n) increasing to ∞ and set $X = \lim X^{U_n}$. This is clearly a global strong solution of equation (5.1.5), and up to indistinguishability the only one. ∎

[19] This is guaranteed by any of the inequalities (5.2.7)–(5.2.11). For (5.2.12) to make sense and to hold, $\boldsymbol{F}$ needs to be defined only on $X, Y \in \mathfrak{S}_{p,M}^{n}$. See also remark 5.2.20.

Exercise 5.2.16 Suppose $\boldsymbol{Z}$ is a quasi-left-continuous L^p-integrator for some $p \ge 2$. Then its previsible controller Λ is continuous and can be chosen strictly increasing; THE time transformation $T^{\bullet}$ is then continuous and strictly increasing as well. Then the Picard iterates $X^{(n)}$ for equation (5.2.18) converge to the solution X in the sense that for all $\lambda < \infty$

$$\Big\| \big|X^{(n)} - X\big|^{\star}_{T^\lambda} \Big\|_{L^p(\mathbb{P})} \xrightarrow[n\to\infty]{} 0 .$$

(i) Conclude from this that if both C and $\boldsymbol{Z}$ are local martingales, then so is X. (ii) Use factorization to extend the previous statement to $p \ge 1$. (iii) Suppose the T^λ are chosen bounded as they can be, and both C and $\boldsymbol{Z}$ are p-integrable martingales. Then X_{T^λ} is a p-integrable martingale on $\{\mathcal{F}_{T^\lambda} : \lambda \ge 0\}$.

Exercise 5.2.17 We say $\boldsymbol{Z}$ is ***surely controlled*** if there exists a right-continuous increasing sure (deterministic) function $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\eta_0 = 0$ so that $d\Lambda^{\langle q \rangle}[\boldsymbol{Z}] \le d\eta$. In this case the stopping times T^λ of equation (5.2.3) are surely bounded from below by the instants

$$t^\lambda \stackrel{\text{def}}{=} \inf\{t : \eta_t \ge \lambda\} \xrightarrow[\lambda\to\infty]{} \infty ,$$

which can be viewed as a deterministic time transform; and if ${}^0C$ is a surely controlled L^q-integrator and ${}^0\boldsymbol{F}$ is Lipschitz and bounded, then the unique solution of theorem 5.2.15 is a surely controlled L^q-integrator as well.

An example of a surely controlled integrator is a Lévy process, in particular a Wiener process. Its previsible controller is of the form $\Lambda_t = C \cdot t$, with $C = \sup_{\rho, p \le q} C_p^{(\rho)(4.6.30)}$. Here $T^\lambda = \lambda / C$.

Exercise 5.2.18 (i) Let W be a standard scalar Wiener process. Exercises 5.2.16 and 5.2.17 together show that the Doléans–Dade exponential of any multiple of W is a martingale. Deduce the following estimates: for $m \in \mathbb{R}$, $p > 1$, and $t \ge 0$

$$\|\mathscr{E}_t[mW]\|_{L^p} = e^{m^2(p-1)t/2} \quad \text{and} \quad \Big\| e^{|mW|^\star_t} \Big\|_{L^p} \le 2p' \cdot e^{m^2 pt/2} ,$$

$p' \stackrel{\text{def}}{=} p/(p-1)$ being the conjugate of p. (ii) Next let $\boldsymbol{Z}_t = (t, W_t^2, \ldots, W_t^d)$, where $\boldsymbol{W}$ is a $d-1$-dimensional standard Wiener process. There are constants B', M' depending only on d, $m \in \mathbb{R}$, $p > 1$, $r > 1$ so that

$$\Big\| |\boldsymbol{Z}^\star|^r_t \cdot e^{m|\boldsymbol{Z}^\star|_t} \Big\|_{L^p} \le B' \cdot t^{r/2} e^{M't} , \tag{5.2.30}$$

$$\|\mathscr{E}_t[mW]\|_{L^p} = e^{m^2(p-1)t/2} , \text{ and } \Big\| e^{|mW|^\star_t} \Big\|_{L^p} \le 2p' \cdot e^{m^2 pt/2} .$$

Exercise 5.2.19 The autologous coefficients f_η of example 5.2.7 form a ***locally Lipschitz*** family if (5.2.17) is satisfied at least on bounded paths: for every $n \in \mathbb{N}$ there exists a constant L_n so that

$$\big| f_\eta[x_\bullet] - f_\eta[y_\bullet] \big|_{\bullet-} \le L_n \cdot |x_\bullet - y_\bullet|^\star$$

whenever the paths $x_\bullet, y_\bullet$ satisfy[4] $|x|^\star_\infty \le n$ and $|y|^\star_\infty \le n$. For instance, markovian coupling coefficients that are continuously differentiable evidently form a locally Lipschitz family. Given such f_η, set ${}^n f_\eta[x] \stackrel{\text{def}}{=} f_\eta[{}^n x]$, where ${}^n x$ is the path x stopped just before the first time T^n its length exceeds n: ${}^n x \stackrel{\text{def}}{=} (x - \Delta_{T^n} x)^{T^n}$. Let ${}^n X$ denote the unique solution of the Lipschitz system

$$X = C + {}^n f_\eta[X]_{\bullet-} {*} Z^\eta \quad \text{and set} \quad {}^n T \stackrel{\text{def}}{=} \inf\{t : {}^l X_t \ge n\}$$

for $l > n$. On $[\![0, {}^n T)\!)$, ${}^l X$ and ${}^n X$ agree. Set $\zeta \stackrel{\text{def}}{=} \sup {}^n T$. There is a unique limit X of ${}^n X$ on $[\![0, \zeta)\!)$. It solves $X = C + f_\eta[X]_{\bullet-} {*} Z^\eta$ there, and ζ is its lifetime.

Stability

The solution of the stochastic differential equation (5.2.31) depends of course on the initial condition C, the coupling coefficient $\boldsymbol{F}$, and on the driver $\boldsymbol{Z}$. How?

We follow the lead provided by item 5.1.6 and consider the difference

$$\Delta \overset{\text{def}}{=} X' - X$$

of the solutions to the two equations

$$X = C + F_\eta[X]_{.-}*Z^\eta \tag{5.2.31}$$

and $$X' = C' + F'_\eta[X']_{.-}*Z'^\eta \; .$$

Δ satisfies itself a stochastic differential equation, namely,

$$\Delta = D + G_\eta[\Delta]_{.-}*Z'^\eta \; , \tag{5.2.32}$$

with initial condition

$$\begin{aligned} D &= (C' - C) + \big(F'_\eta[X]_{.-}*Z'^\eta - F_\eta[X]_{.-}*Z^\eta\big) \\ &= (C' - C) + (F'_\eta[X] - F_\eta[X])_{.-}*Z'^\eta + F_\eta[X]_{.-}*(Z'^\eta - Z^\eta) \end{aligned}$$

and coupling coefficients

$$\Delta \mapsto G_\eta[\Delta] \overset{\text{def}}{=} F'_\eta[\Delta + X] - F_\eta[X] \; .$$

To answer our question, how?, we study the size of the difference Δ in terms of the differences of the initial conditions, the coupling coefficients, and the drivers. This is rather easy in the following frequent situation: both $\boldsymbol{Z}$ and $\boldsymbol{Z}'$ are local L^q-integrators for some[12] $q \geq 2$, and the seminorms $\|\ \|^\star_{p,M}$ are defined via (5.2.3) and (5.2.4) from a previsible controller Λ common[20] to both $\boldsymbol{Z}$ and $\boldsymbol{Z}'$; and for some fixed $\gamma < 1$ and $M \geq M^{(5.2.26)}_{L:\gamma}$, $\boldsymbol{F}'$ and with it $\boldsymbol{G}$ satisfies the Lipschitz condition (5.2.12) on page 286, with constant L. In this situation inequality (5.2.23) on page 290 immediately gives the following estimate of Δ:

$$\begin{aligned} \|\Delta\|^\star_{p,M} &\leq \frac{1}{1-\gamma} \cdot \big\|(C'-C)+\big(\boldsymbol{F}'[X]_{.-}*\boldsymbol{Z}'-\boldsymbol{F}[X]_{.-}*\boldsymbol{Z}\big)\big\|^\star_{p,M} \\ &= \frac{1}{1-\gamma} \cdot \big\|(C'-C)+(\boldsymbol{F}'[X]-\boldsymbol{F}[X])_{.-}*\boldsymbol{Z}'+\boldsymbol{F}[X]_{.-}*(\boldsymbol{Z}'-\boldsymbol{Z})\big\|^\star_{p,M}, \end{aligned}$$

which with (5.2.27) implies

$$\begin{aligned} \|X'-X\|^\star_{p,M} \leq \frac{1}{1-\gamma} \cdot \Big(\big\|(C'-C)\big\|^\star_{p,M} &+ \frac{\gamma}{L} \cdot \big\| |\boldsymbol{F}'[X]-\boldsymbol{F}[X]|_\infty \big\|_{p,M} \\ &+ \big\|\boldsymbol{F}[X]_{.-}*(\boldsymbol{Z}'-\boldsymbol{Z})\big\|^\star_{p,M}\Big) \; . \end{aligned} \tag{5.2.33}$$

[20] Take, for instance, for Λ the sum of the canonical controllers of the two integrators.

This inequality exhibits plainly how the solution X depends on the ingredients $C, \boldsymbol{F}, \boldsymbol{Z}$ of the stochastic differential equation (5.2.31). We shall make repeated use of it, to produce an algorithm for the pathwise solution of (5.2.31) in section 5.4, and to study the differentiability of the solution in parameters in section 5.3. Very frequently only the initial condition and coupling coefficient are perturbed, $\boldsymbol{Z} = \boldsymbol{Z}'$ staying unaltered. In this special case (5.2.33) becomes

$$\left\| X' - X \right\|^{\star}_{p,M} \le \frac{1}{1-\gamma} \cdot \left(\left\| (C' - C) \right\|^{\star}_{p,M} + \frac{\gamma}{L} \cdot \left\| \left| \boldsymbol{F}'[X] - \boldsymbol{F}[X] \right|_{\infty} \right\|_{p,M} \right) \tag{5.2.34}$$

or, with the roles of X, X' reversed,

$$\left\| X' - X \right\|^{\star}_{p,M} \le \frac{1}{1-\gamma} \cdot \left(\left\| (C' - C) \right\|^{\star}_{p,M} + \frac{\gamma}{L} \cdot \left\| \left| \boldsymbol{F}'[X'] - \boldsymbol{F}[X'] \right|_{\infty} \right\|_{p,M} \right) . \tag{5.2.35}$$

Remark 5.2.20 In the case $\boldsymbol{F} = \boldsymbol{F}'$ and $\boldsymbol{Z} = \boldsymbol{Z}'$, (5.2.34) boils down to

$$\left\| X' - X \right\|^{\star}_{p,M} \le \frac{1}{1-\gamma} \cdot \left\| (C' - C) \right\|^{\star}_{p,M} . \tag{5.2.36}$$

Assume for example that $\boldsymbol{Z}, \underline{\boldsymbol{Z}}$ are two L^q-integrators and Λ a common controller.[20] Consider the equations $F = X + H^\eta[F]_{\leftharpoondown} * \underline{Z}^\eta$, $\eta = 1, \ldots d$, H^η Lipschitz. The map that associates with X the unique solution $F_\eta[X]$ is according to (5.2.36) a Lipschitz coupling coefficient in the weak sense of inequality (5.2.12) on page 286. To paraphrase: "the solution of a Lipschitz stochastic differential equation is a Lipschitz coupling coefficient in its initial value and as such can function in another stochastic differential equation." In fact, this Picard-norm Lipschitz coupling coefficient is even differentiable, provided the H^η are (exercise 5.3.7).

Exercise 5.2.21 If $\boldsymbol{F}[0] = 0$, then $\left\| C \right\|^{\star}_{p,M} \le 2 \cdot \left\| X \right\|^{\star}_{p,M}$.

Lipschitz and Pathwise Continuous Versions Consider the situation that the initial condition C and the coupling coefficient $\boldsymbol{F}$ depend on a parameter u that ranges over an open subset U of some seminormed space $(E, \| \ \|_E)$. Then the solution of equation (5.2.18) will depend on u as well: in obvious notation

$$X[u] = C[u] + \boldsymbol{F}\big[u, X[u]\big]_{\leftharpoondown} * \boldsymbol{Z} . \tag{5.2.37}$$

A cheap consequence of the stability results above is the following observation, which is used on several occasions in the sequel. Suppose that the initial condition and coupling coefficient in (5.2.37) are jointly Lipschitz, in the sense that there exists a constant L such that for all $u, v \in U$ and all $X \in \mathfrak{S}^{\star n}_{p,M}$ (where $2 \le p \le q$ and $M > M^{(5.2.20)}_{p,L}$)

$$\left\| C[v] - C[u] \right\|^{\star}_{p,M} \le L \cdot \| v - u \|_E \tag{5.2.38}$$

and $$\left\| \left| \boldsymbol{F}[v, Y] - \boldsymbol{F}[u, X] \right|_{\infty} \right\|^{\star}_{p,M} \le L \cdot \left(\| v - u \|_E + \left\| Y - X \right\|^{\star}_{p,M} \right) ;$$

hence $$\left\| \left| \boldsymbol{F}[v, X] - \boldsymbol{F}[u, X] \right|_{\infty} \right\|^{\star}_{p,M} \le L \cdot \| v - u \|_E . \tag{5.2.39}$$

Then inequality (5.2.34) implies the Lipschitz dependence of $X[u]$ on $u \in U$:

Proposition 5.2.22 *In the presence of (5.2.38) and (5.2.39) we have for all $u, v \in U$*

$$\left\|X[v] - X[u]\right\|_{p,M}^{\star} \le \frac{L+\gamma}{1-\gamma} \cdot \|v-u\|_E . \tag{5.2.40}$$

Corollary 5.2.23 *Assume that the parameter domain U is finite-dimensional, $\boldsymbol{Z}$ is a local L^p-integrator for some p strictly larger than $\dim U$, and for all $u, v \in U$*[4]

$$\left\|X[v] - X[u]\right\|_{p,M}^{\star} \le const \cdot |v-u| . \tag{5.2.41}$$

Then the solution processes $X_{\boldsymbol{.}}[u]$ can be chosen in such a way that for every $\omega \in \Omega$ the map $u \mapsto X_{\boldsymbol{.}}[u](\omega)$ from U to $\mathscr{D}^n$ is continuous.[21]

Proof. For fixed t and λ the Lipschitz condition (5.2.41) gives

$$\Big\| \sup_{s \le T^\lambda} \big\|X[v] - X[u]\big\|_s \Big\|_{L^p} \le const \cdot e^{M\lambda} \cdot |v-u| ,$$

which implies $\quad \mathbb{E}\Big[\sup_{s\le t} \big|X[v]-X[u]\big|_s^p \cdot [t < T^\lambda]\Big] \le const \cdot |v-u|^p .$

According to Kolmogorov's lemma A.2.37 on page 384, there is a negligible subset N_λ of $[t < T^\lambda] \in \mathcal{F}_t$ outside which the maps $u \mapsto X[u]^t_{\boldsymbol{.}}(\omega)$ from U to paths in $\mathscr{D}^n$ stopped at t are, after modification, all continuous in the topology of uniform convergence. We let λ run through a sequence (λ_n) that increases without bound. We then either throw away the nearly empty set $\bigcup_n N_{\lambda^{(n)}}$ or set $X[u]_{\boldsymbol{.}} = 0$ there. ∎

In particular, when $\boldsymbol{Z}$ is continuous it is a local L^q-integrator for all $q < \infty$, and $u \mapsto X_{\boldsymbol{.}}[u](\omega)$ can be had continuous for every $\omega \in \Omega$.

If $\boldsymbol{Z}$ is merely an L^0-integrator, then a change of measure as in the proof of theorem 5.2.15 allows the same conclusion, except that we need Lipschitz conditions that do not change with the measure:

Theorem 5.2.24 *In (5.2.37) assume that C and F are Lipschitz in the finite-dimensional parameter u in the sense that for all $u, v \in U$ and all $X \in \mathfrak{D}^n$*

both[4] $$\big|C[v] - C[u]\big|^{\star} \le L \cdot |v-u|$$

and $$\sup_\eta \big|F_\eta[v,X] - F_\eta[u,X]\big| \le L \cdot |v-u| ,$$

nearly. Then the solutions $X_{\boldsymbol{.}}[u]$ of (5.2.37) can be chosen in such a way that for every $\omega \in \Omega$ the map $u \mapsto X_{\boldsymbol{.}}[u](\omega)$ from U to $\mathscr{D}^n$ is continuous.[21] ∎

[21] The pertinent topolgy on path spaces is the topology of uniform convergence on compacta.

Differential Equations Driven by Random Measures

By definition 3.10.1 on page 173, a random measure ζ is but an "$\boldsymbol{H}$-tuple of integrators," $\boldsymbol{H}$ being the auxiliary space. Instead of $\int \sum_\eta F_{\eta s} dZ_s^\eta$ or $\int \int_\eta F_{\eta s} dZ_s^{d\eta}$ we write $\int F(\eta, s)\, \zeta(d\eta, ds)$, but that is in spirit the sum total of the difference. Looking at a random measure this way, as "a long vector of tiny integrators" as it were, has already paid nicely (see theorem 4.3.24 and theorem 4.5.25). We shall now see that stochastic differential equations driven by random measures can be handled just like the ones driven by slews of integrators that we have treated so far. In fact, the following is but a somewhat repetitious reprise of the arguments developed above. It simplifies things a little to ***assume that ζ is spatially bounded.***

From this point of view the stochastic differential equation driven by ζ is

$$X_t = C_t + \int_0^t F[\eta, X_{\bullet}]_{s-}\, \zeta(d\eta, ds)$$

or, equivalently, $$X = C + F[\,\bullet\,, X_{\bullet}]_{\bullet -} {*} \zeta\,, \tag{5.2.42}$$

with $$F : \boldsymbol{H} \times \mathfrak{D}^n \to \mathfrak{D}^n \qquad \text{suitable.}$$

We expect to solve (5.2.42) under the strong Lipschitz condition (5.2.8) on page 286, which reads here as follows: at any stopping time T

$$\big|\big(F[\,\bullet\,, Y] - F[\,\bullet\,, X]\big)_T\big|_{\infty p} \le L \cdot \big|\big(Y - X\big)_t^\star\big|_p \quad \text{a.s.,}$$

where $\big|F[\,\bullet\,, X]_s\big|_\infty$ is the n-vector $\big(\sup \big|F^\nu[\eta, X]_s\big| : \eta \in \boldsymbol{H}\big)_{\nu=1}^n$,

$\big|F[\,\bullet\,, X]_{\bullet}^\star\big|_\infty$ its (vector-valued) maximal process,

and $\big|F[\,\bullet\,, X]_{\bullet}^\star\big|_{\infty p} \stackrel{\text{def}}{=} \Big(\sum_\nu \big|F^\nu[\,\bullet\,, X]_{\bullet}^\star\big|_\infty^p\Big)^{1/p}$ the length of the latter.

As a matter of fact, if ζ is an L^q-integrator and $2 \le p \le q$, then the following rather much weaker "p-mean Lipschitz condition," analog of inequality (5.2.10), should suffice: for any $X, Y \in \mathfrak{D}^n$ and any predictable stopping time T

$$\Big\| \big|F[\,\bullet\,, Y]_{T-} - F[\,\bullet\,, X]_{T-}\big|_{\infty p} \Big\|_{L^p} \le L \cdot \Big\| \big|\big(Y - X\big)_{T-}^\star\big|_p \Big\|_{L^p} .$$

Assume this then and let $\Lambda = \Lambda^{\langle q \rangle}[\zeta]$ be the previsible controller provided by theorem 4.5.25 on page 251. With it goes THE time transformation (5.2.3), and with the help of the latter we define the seminorms $\|\ \|_{p,M}$ and $\|\ \|_{p,M}^\star$ as in definition (5.2.4) on page 283. It is a simple matter of shifting the η from a subscript on $\boldsymbol{F}$ to an argument in F, to see that lemma 5.2.1 persists. Using definition (5.2.26) on page 290, we have for any $\gamma \in (0,1)$

$$\big\| F_{\bullet -} {*} \zeta \big\|_{p,M}^\star \le \frac{\gamma}{L} \cdot \big\| |F|_\infty \big\|_{p,M}^\star$$

and $$\left\| \mathfrak{U}[Y] - \mathfrak{U}[X] \right\|^{\star}_{p,M} \le \gamma \cdot \left\| Y - X \right\|^{\star}_{p,M} ,$$

where of course $$\mathfrak{U}[X]_t \stackrel{\text{def}}{=} C_t + \int_0^t F[\eta, X]_{s-}\, \zeta(d\eta, ds)$$

and $$M > M_{L:\gamma} \stackrel{\text{def}}{=} (10qL/\gamma)^q \vee 1 .$$

We see that *as long as* $\mathfrak{S}^{\star n}_{p,M}$ *contains* ${}^0C \stackrel{\text{def}}{=} C + F[\,\cdot\,, 0]_{.-}{*}\zeta$ *it contains a unique solution of equation* (5.2.42). If ζ is merely an L^0-random measure, then we reduce as in theorem 5.2.15 the situation to the previous one by invoking a factorization, this time using corollary 4.1.14 on page 208:

Proposition 5.2.25 *Equation (5.2.42) has a unique global solution.*

The stability inequality (5.2.34) for the difference $\Delta \stackrel{\text{def}}{=} X' - X$ of the solutions of two stochastic differential equations

$$X' = C' + F'[\,\cdot\,, X']_{.-}{*}\zeta'$$

and $$X \;= C \;+\; F[\,\cdot\,, X]_{.-}{*}\zeta \;,$$

which satisfies $$\Delta \;= D \;+\; G[\,\cdot\,, \Delta]_{.-}{*}\zeta'$$

with $$\begin{aligned} D = (C' - C) &+ (F'[\,\cdot\,, X] - F[\,\cdot\,, X])_{.-}{*}\zeta' \\ &+\; F[\,\cdot\,, X]_{.-}{*}(\zeta' - \zeta) \end{aligned}$$

and $$G[\,\cdot\,, \Delta] = F'[\,\cdot\,, \Delta + X] - F'[\,\cdot\,, X] \;,$$

persists *mutatis mutandis*. Assuming that both F and F' are Lipschitz with constant L, that $\|\;\|^{\star}_{p,M}$ is defined from a previsible controller Λ common to both ζ and ζ',[22] and that M has been chosen strictly larger[23] than $M^{\diamond(5.2.20)}_{p,L}$, the analog of inequality (5.2.23) results in these estimates of Δ:

$$\begin{aligned} \left\|\Delta\right\|^{\star}_{p,M} &\le \frac{1}{1-\gamma} \cdot \left\| (C' - C) + \big(F'[\,\cdot\,, X]_{.-}{*}\zeta' - F[\,\cdot\,, X]_{.-}{*}\zeta\big) \right\|^{\star}_{p,M} \\ &\le \frac{1}{1-\gamma} \cdot \Big(\left\|(C' - C)\right\|^{\star}_{p,M} + \frac{\gamma}{L} \cdot \left\| \left| F'[\,\cdot\,, X] - F[\,\cdot\,, X] \right|_{\infty} \right\|_{p,M} \\ &\qquad\qquad + \left\| F[\,\cdot\,, X]_{.-}{*}(\zeta' - \zeta) \right\|^{\star}_{p,M} \Big) \;. \end{aligned}$$

This inequality shows neatly how the solution X depends on the ingredients C, F, ζ of the equation. Additional assumptions, such as that $\zeta' = \zeta$ or that both $\zeta = \zeta'$ and $C' = C$ simplify these inequalities in the manner of the inequalities (5.2.34)–(5.2.36) on page 294.

---

[22] Take, for instance, $\Lambda_t \stackrel{\text{def}}{=} \Lambda^{\langle q \rangle}_t[\zeta] + \Lambda^{\langle q \rangle}_t[\zeta'] + \epsilon t ,\ 0 < \epsilon \ll 1 .$

[23] The point being that p and M must be chosen so that $\mathfrak{U}$ is strictly contractive on $\mathfrak{S}^{\star n}_{p,M}$.

The Classical SDE

The "classical" stochastic differential equation is the markovian equation

$$X = C + \boldsymbol{f}(X)*\boldsymbol{Z} \quad\text{or}\quad X_t = C + \int_0^t f_\eta(X)_s \, dZ_s^\eta \;, \qquad t \ge 0 \;,$$

where the initial condition C is constant in time, $\boldsymbol{f} = (f_0, f_1, \ldots, f_d)$ are (at least measurable) vector fields on the state space $\mathbb{R}^n$ of X, and the driver $\boldsymbol{Z}$ is the $d{+}1$-tuple $\boldsymbol{Z}_t = (t, W_t^1, \ldots, W_t^d)$, $\boldsymbol{W}$ being a standard Wiener process on a filtration to which both $\boldsymbol{W}$ and the solution are adapted. The classical SDE thus takes the form

$$X_t = C + \int_0^t f_0(X_s)\, ds + \textstyle\sum_{\eta=1}^d \displaystyle\int_0^t f_\eta(X_s)\, dW_s^\eta \;. \tag{5.2.43}$$

In this case the controller is simply $\Lambda_t = d \cdot t$ (exercise 4.5.19) and thus THE time transformation is simply $T^\lambda = \lambda/d$. The Picard norms of a process X become simply

$$\|X\|_{p,M}^\star = \sup_{t>0} e^{-Mdt} \cdot \Big\| \big|X_t^\star\big|_p \Big\|_{L^p(\mathbb{P})}^* \;.$$

If the coupling coefficient $\boldsymbol{f}$ is Lipschitz, then the solution of (5.2.43) grows at most exponentially in the sense that $\| \big|X_t^\star\big|_p \|_{L^p(\mathbb{P})} \le Const \cdot e^{Mdt}$. The stability estimates, etc., translate similarly.

5.3 Stability: Differentiability in Parameters

We consider here the situation that the initial condition C and the coupling coefficient $\boldsymbol{F}$ depend on a parameter u that ranges over an open subset U of some seminormed space $(E, \| \;\|_E)$. Then the solution of equation (5.2.18) will depend on u as well: in obvious notation

$$X[u] = C[u] + \boldsymbol{F}\big[u, X[u]\big]_{.-} * \boldsymbol{Z} \;. \tag{5.3.1}$$

We have seen in item 5.1.7 that in the case of an ordinary differential equation the solution depends differentiably on the initial condition and the coupling coefficient. This encourages the hope that our $X[u]$, too, will depend differentiably on u when both $C[u]$ and $\boldsymbol{F}[u, \cdot\,]$ do. This is true, and the goal of this section is to prove several versions of this fact.

Throughout the section ***the minimal assumptions (i)–(iii)*** of page 272 are in effect. In addition we will require that $\boldsymbol{Z} = (Z^1, \ldots, Z^d)$ is a local $L^q(\mathbb{P})$-integrator for some[12] ***q strictly larger than 2*** – except when this is explicitely rescinded on occasion. This requirement provides us with the previsible controller $\Lambda^{\langle q \rangle}[\boldsymbol{Z}]$, with THE time transformation (5.2.3), and the Picard norms[11] $\|\; \|_{p,M}^\star$ of (5.2.4). We also have settled on a modulus of contractivity $\gamma \in (0,1)$ to our liking. The coupling coefficients $\boldsymbol{F}[u,\cdot\,]$ are assumed to be Lipschitz in the sense of inequality (5.2.12) on page 286:

$$\Big\| \big|\boldsymbol{F}[u,Y] - \boldsymbol{F}[u,X]\big|_\infty \Big\|_{p,M} \le L \cdot \big\| X - Y \big\|^\star_{p,M} , \tag{5.3.2}$$

with ***Lipschitz constant*** L ***independent*** of the parameter u and of

$$p \in (2,q] , \quad \text{and} \quad M \ge M^{(5.2.26)}_{L:\gamma} . \tag{5.3.3}$$

Then any stochastic differential equation driven by $\boldsymbol{Z}$ and satisfying the Picard-norm Lipschitz condition (5.3.2) and equation (5.2.28) has its solution in $\mathfrak{S}^\star_{p,M}$, whatever such p and M. In particular, $X[u] \in \mathfrak{S}^{\star n}_{p,M}$ for all $u \in U$ and all (p,M), as in (5.3.3).

For the notation and terminology concerning differentiation refer to definitions A.2.45 on page 388 and A.2.49 on page 390.

Example 5.3.1 Consider first the case that U is an open convex subset of $\mathbb{R}^k$ and the coupling coefficient $\boldsymbol{F}$ of (5.3.1) is markovian (see example 5.2.4): $F_\eta[u,X] = f_\eta(u,X)$. Suppose the $f_\eta : U \times \mathbb{R}^n \to \mathbb{R}^n$ have a continuous bounded derivative $Df_\eta = (D_1 f_\eta, D_2 f_\eta)$. Then just as in example A.2.48 on page 389, F_η is not necessarily Fréchet differentiable as a map from $U \times \mathfrak{S}^{\star n}_{p,M}$ to $\mathfrak{S}^{\star n}_{p,M}$. It is, however, weakly uniformly differentiable, and the partial $D_2 F_\eta[u,X]$ is the continuous linear operator from $\mathfrak{S}^{\star n}_{p,M}$ to itself that operates on a $\xi \in \mathfrak{S}^{\star n}_{p,M}$ by applying the $n{\times}n$-matrix $D_2 f_\eta(u,X)$ to the vector $\xi \in \mathbb{R}^n$: for $\varpi \in \boldsymbol{B}$

$$\big(D_2 F_\eta[u, X(\varpi)]\cdot\xi\big)(\varpi) = D_2 f_\eta\big(u, X(\varpi)\big)\cdot\xi(\varpi) .$$

The operator norm of $D_2 F_\eta[u,X]$ is bounded by $\sup_{u,x} |Df_\eta(u,x)|_1$, independently of $u \in U$ and p, where $|D|_1 \stackrel{\text{def}}{=} \sum_{\nu,\kappa} |D^\nu_\kappa|$ on matrices D.

Example 5.3.2 The previous example has an extension to autologous coupling coefficients. Suppose the adapted map[17] $f : U \times \mathscr{D}^n \to \mathscr{D}^n$ has a continuous bounded Fréchet derivative. Let us make this precise: at any point $(u, x_.)$ in $U \times \mathscr{D}^n$ there exists a linear map $Df[u,x_.] : E \times \mathscr{D}^n \to \mathscr{D}^n$ such that for all t

$$\big|Df[u,x_.]\big|^\star_t \stackrel{\text{def}}{=} \sup\left\{ \left| Df[u,x_.]\cdot \begin{pmatrix} \xi \\ \Xi_. \end{pmatrix} \right|^\star_t : \|\xi\|_E + |\Xi|^\star_t \le 1 \right\} \le L < \infty$$

and $\quad \big|Df[v,y_.] - Df[u,x_.]\big|^\star_t \to 0 \quad$ as $\quad \|v-u\|_E + |y_. - x_.|^\star_t \to 0$,

with L independent of $(u, x_.)$, and such that

$$Rf[u,x_.;v,y_.] \stackrel{\text{def}}{=} f[v,y_.] - f[u,x_.] - Df[u,x_.]\cdot \begin{pmatrix} v-u \\ y_. - x_. \end{pmatrix} \quad \text{has}$$

$$|Rf[u,x_.;v,y_.]|^\star_t = o\big(\|v-u\|_E + |y_. - x_.|^\star_t\big) .$$

According to Taylor's formula of order one (see lemma A.2.42), we have[24]

$$Rf[u,x_.;v,y_.]=\int_0^1\Big(Df[u^\sigma,x^\sigma_.]-Df[u,x_.]\Big)\,d\sigma\cdot\begin{pmatrix}v-u\\ y_.-x_.\end{pmatrix}\;,$$

where $u^\sigma \overset{\text{def}}{=} u+\sigma(v-u)$ and $x^\sigma_. \overset{\text{def}}{=} x_.+\sigma(y_.-x_.)$.

Now the coupling coefficient F corresponding to f is defined at $X\in\mathfrak{D}^n$ by

$$F[u,X]_.(\omega)\overset{\text{def}}{=} f[u,X_.(\omega)]$$

and has weak derivative $DF[u,X]:\begin{pmatrix}\xi\\ \Xi\end{pmatrix}\mapsto Df[u,X]\cdot\begin{pmatrix}\xi\\ \Xi_.\end{pmatrix}\;,\quad \Xi\in\mathfrak{S}^{\star n}_{p,M}$.

Indeed, for $1/p^\circ=1/p+1/r$ and $M^\circ=M+R$,

$$\begin{aligned}\big\|RF[u,X;v,Y]\big\|^\star_{p^\circ,M^\circ}&\le\int_0^1\big\|Df[u^\sigma,X^\sigma]-Df[u,X]\big\|^\star_{r,R}\,d\sigma\\ &\qquad\times\Big(\|v-u\|_E+\big\|Y-X\big\|^\star_{p,M}\Big)\\ &=o\Big(\|v-u\|_E+\big\|Y-X\big\|^\star_{p,M}\Big)\end{aligned}$$

on the grounds that $\big|Df[u^\sigma,X^\sigma]-Df[u,X]\big|^\star_{T^\lambda-}\to 0$

as $\|v-u\|_E+\big\|Y-X\big\|^\star_{p,M}\to 0$, pointwise and boundedly for every λ, σ, and $\omega\in\Omega$: the weak uniform differentiability follows. ∎

Exercise 5.3.3 Extend this to randomly autologous coefficients $f[\omega,u,x_.]$: Assume that the family $\{f[\omega,u,x_.]:\omega\in\Omega\}$ of autologous coefficients is equidifferentiable and has $\sup_\omega\big|Df[\omega,u,x_.]\big|^\star_t\le L$. Show that then $F[u,X]_.(\omega)\overset{\text{def}}{=} f[\omega,u,X_.(\omega)]$ is weakly differentiable on $\mathfrak{S}^{\star n}_{p,M}$ with

$$DF[u,X]:\begin{pmatrix}\xi\\ \Xi\end{pmatrix}\mapsto Df[\omega,u,X(\omega)]\cdot\begin{pmatrix}\xi\\ \Xi_.(\omega)\end{pmatrix}\;.$$

Example 5.3.4 Example 5.2.8 exhibits a coupling coefficient $\mathfrak{S}^{\star n}_{p,M}\to\mathfrak{S}^{\star n}_{p,M}$ that is linear and bounded for all (p,M) and *ipso facto* uniformly Fréchet differentiable. We leave to the reader the chore of finding the conditions that make examples 5.2.10–5.2.11 weakly differentiable. Suppose the maps $(u,X)\mapsto F[u,X],G[u,X]$ are both weakly differentiable as functions from $U\times\mathfrak{S}^{\star n}_{p,M}$ to $\mathfrak{S}^{\star n}_{p,M}$. Then so is $(u,X)\mapsto F\big[u,G[u,X]\big]$.

Example 5.3.5 Let $\mathcal{T}$ be an adapted partition of $[0,\infty)\times\mathscr{D}^n$. The scalæfication map $x_.\mapsto x^{\mathcal{T}}_.$ is a differentiable autologous Lipschitz coupling coefficient. If F is another, then the coupling coefficient F' of equation (5.4.7) on page 312, defined by $F':Y\mapsto F[Y^{\mathcal{T}}]^{\mathcal{T}}$, is yet another.

[24] To see this, reduce to the scalar case by applying a continuous linear functional.

The Derivative of the Solution

Let us then stipulate for the remainder of this section that in the stochastic differential equation (5.3.1) the initial condition $u \mapsto C[u]$ is weakly differentiable as a map from U to $\mathfrak{S}^{\star n}_{p,M}$, and that the coupling coefficients

$$F_\eta : U \times \mathfrak{S}^{\star n}_{p,M} \to \mathfrak{S}^{n}_{p,M} \ , \ F_\eta : (u, X) \mapsto F_\eta[u, X] \ ,$$

are weakly differentiable as well. What is needed at this point is a candidate $DX[u]$ for the derivative of $v \mapsto X[v]$ at $u \in U$. In order to get an idea what it might be, let us assume that X is in fact weakly differentiable. With

$$v = u + \tau\xi \ , \ \ \|\xi\|_E = 1 \ , \ \ \tau = \|v - u\|_E \ ,$$

this can be written as

$$X[v] = X[u] + \tau DX[u]{\cdot}\xi + RX[u; v] \ , \tag{5.3.4}$$

where $\quad \big\|RX[u; v]\big\|_{p^\circ,M^\circ} = o(\tau) \quad$ for $p^\circ{<}p \ , \ M^\circ{>}M$.

On the other hand,

$$\begin{aligned} X[v] &= X[u] + C[v] - C[u] + \Big\{\boldsymbol{F}\big[v, X[v]\big] - \boldsymbol{F}\big[u, X[u]\big]\Big\}_{\bullet-} {*}\boldsymbol{Z} \\ &= X[u] + DC[u]{\cdot}(v-u) + RC[u; v] \\ &\qquad + \left\{D\boldsymbol{F}\big[u, X[u]\big]{\cdot}\left(\begin{matrix} v-u \\ X[v] - X[u]\end{matrix}\right)\right\}_{\bullet-} {*}\boldsymbol{Z} \\ &\qquad + R\boldsymbol{F}[u, X[u]; v, X[v]]_{\bullet-} {*}\boldsymbol{Z} \ , \end{aligned}$$

which translates to

$$X[v] = X[u] + \tau DC[u]{\cdot}\xi + \tau\left\{D\boldsymbol{F}\big[u, X[u]\big]{\cdot}\left(\begin{matrix}\xi \\ DX[u]{\cdot}\xi\end{matrix}\right)\right\}_{\bullet-} {*}\boldsymbol{Z} \tag{5.3.5}$$

$$+ RC + R\boldsymbol{F}_{\bullet-} {*}\boldsymbol{Z} + (D_2\boldsymbol{F}{\cdot}RX)_{\bullet-} {*}\boldsymbol{Z} \ . \tag{5.3.6}$$

In the previous line the arguments $[\,\cdot\,]$ of RC, RX, and $R\boldsymbol{F}$ are omitted from the display. In view of inequality (5.2.5) the entries in line (5.3.6) are $o(\tau)$, as is the last term in (5.3.4). Thus both equations (5.3.4) and (5.3.5) for $X[v]$ are of the form $X[u] + \mathit{const} \cdot \tau + o(\tau)$ when measured in the pertinent Picard norm, which here is $\|\ \|^{\star}_{p^\circ,M^\circ}$. Clearly,[25] then, the coefficient of τ on both sides must be the same. We arrive at

$$\begin{aligned} DX[u]{\cdot}\xi = DC[u]{\cdot}\xi &+ \big(D_1\boldsymbol{F}\big[u, X[u]\big]{\cdot}\xi\big)_{\bullet-} {*}\boldsymbol{Z} \\ &+ \big\{D_2\boldsymbol{F}\big[u, X[u]\big] \cdot (DX[u]{\cdot}\xi)\big\}_{\bullet-} {*}\boldsymbol{Z} \ . \end{aligned} \tag{5.3.7}$$

[25] To enhance the clarity apply a linear functional that is bounded in the pertinent norm, thus reducing this to the case of real-valued functions. The Hahn–Banach theorem A.2.25 provides the generality stated.

For every fixed $u \in U$ and $\xi \in E$ we see here a linear stochastic differential equation for a process[2] $DX[u]{\cdot}\xi \in \mathfrak{D}^n$, whose initial condition is the sum $DC[u]{\cdot}\xi + \big(D_1\boldsymbol{F}\big[u, X[u]\big]{\cdot}\xi\big)_{\!.-}{*}\boldsymbol{Z}$ and whose coupling coefficient is given by $\Xi \mapsto D_2\boldsymbol{F}\big[u, X[u]\big]{\cdot}\Xi$, which by exercise A.2.50 (c) has a Lipschitz constant less than L. Its solution depends linearly on ξ, lies in $\mathfrak{S}^{\star n}_{p,M}$ for all pairs (p, M) as in (5.3.3), and there it has size (see inequality (5.2.23) on page 290)

$$\big\|DX[u]{\cdot}\xi\big\|^{\star}_{p,M} \le \frac{1}{1-\gamma}\Big(\big\|DC[u]{\cdot}\xi\big\|^{\star}_{p,M} + \frac{\gamma}{L}{\cdot}\big\|D_1\boldsymbol{F}\big[u, X[u]\big]{\cdot}\xi\big\|^{\star}_{p,M}\Big)$$
$$\le \mathit{const} \cdot \|\xi\|_E < \infty\ .$$

Thus if $X[\,\cdot\,]$ is differentiable, then its derivative $DX[u]$ must be given by (5.3.7). Therefore $DX[u]$ is henceforth *defined* as the linear map from E to $\mathfrak{S}^{\star}_{p,M}$ that sends ξ to the unique solution of (5.3.7). It is necessarily our candidate for the derivative of $X[\,\cdot\,]$ at u, and it does in fact do the job:

Theorem 5.3.6 *If $v \mapsto C[v]$ and the $(v, X) \mapsto F_\eta[v, X]$, $\eta = 1, \ldots, d$, are weakly (equi)differentiable maps into $\mathfrak{S}^{\star n}_{p,M}$, then $v \mapsto X[v] \in \mathfrak{S}^{\star n}_{p,M}$ is weakly (equi)differentiable, and its derivative at a point $u \in U$ is the linear map $DX[u]$ defined by (5.3.7).*

Proof. It has to be shown that, for every $p^\circ \in [2, p)$ and $M^\circ > M$,

$$RX[u; v] \stackrel{\text{def}}{=} X[v] - X[u] - DX[u]{\cdot}(v{-}u)$$

has
$$\big\|RX[u; v]\big\|^{\star}_{p^\circ,M^\circ} = o(\|v{-}u\|_E)\ .$$

Now it is a simple matter of comparing equalities (5.3.4) and (5.3.5) to see that, in view of (5.3.7), $RX[u; v]$ satisfies the stochastic differential equation

$$RX[u; v] = RC[u; v] + R\boldsymbol{F}[u, X[u]; v, X[v]]_{.-}{*}\boldsymbol{Z}$$
$$+ \big\{D_2\boldsymbol{F}\big[u, X[u]\big]{\cdot}RX[u; v]\big\}_{.-}{*}\boldsymbol{Z}\ ,$$

whose Lipschitz constant is L. According to (5.2.23) on page 290, therefore,

$$\big\|RX[u; v]\big\|^{\star}_{p^\circ,M^\circ} \le \frac{1}{1{-}\gamma} \cdot \big\|RC[u; v] + R\boldsymbol{F}[u, X[u]; v, X[v]]_{.-}{*}\boldsymbol{Z}\big\|^{\star}_{p^\circ,M^\circ}\ .$$

Since $\big\|RC[u; v]\big\|_{p^\circ,M^\circ} = o\big(\|v{-}u\|_E\big)$ as $v \to u$, all that needs showing is

$$\big\|R\boldsymbol{F}[u, X[u]; v, X[v]]_{.-}{*}\boldsymbol{Z}\big\|^{\star}_{p^\circ,M^\circ} = o\big(\|v{-}u\|_E\big) \ \text{ as } \ v \to u\ ,$$

and this follows via inequality (5.2.5) on page 284 from

$$\big\|RF_\eta[u, X[u]; v, X[v]]\big\|_{p^\circ,M^\circ} = o\big(\|v{-}u\|_E + \big\|X[v] - X[u]\big\|^{\star}_{p,M}\big)$$

by A.2.50 (c) and (5.2.40):
$$= o\big(\|v{-}u\|_E\big) \ \text{ as } \ v \to u\ , \quad \eta = 1, \ldots d\ .$$

If $C, \boldsymbol{F}$ are weakly uniformly differentiable, then the estimates above are independent of u, and $X[u]$ is in fact uniformly differentiable in u. ❏

Exercise 5.3.7 Taking $E \stackrel{\text{def}}{=} \mathfrak{S}^{\star n}_{p,M}$, show that the coupling coefficient of remark 5.2.20 is differentiable.

Pathwise Differentiability

Consider now the difference $\boldsymbol{D} \stackrel{\text{def}}{=} DX[v] - DX[u]$ of the derivatives at two different points u, v of the parameter domain U, applied to an element ξ of E_1.[26] According to equation (5.3.7) and inequality (5.2.34), $\boldsymbol{D}$ satisfies the estimate

$$\left\|\boldsymbol{D}\cdot\xi\right\|^{\star}_{p,M} \le \frac{1}{1-\gamma} \cdot \Big(\left\|(DC[v] - DC[u])\cdot\xi\right\|^{\star}_{p,M} + \frac{\gamma}{L} \cdot \left\| \left|(D\boldsymbol{F}[v, X[v]] - D\boldsymbol{F}\big[u, X[u]\big])\cdot\xi\right|_{\infty} \right\|^{\star}_{p,M} + \frac{\gamma}{L} \cdot \left\| \big(D_2\boldsymbol{F}[v, X[v]] - D_2\boldsymbol{F}\big[u, X[u]\big]\big) \cdot DX[u]\cdot\xi \right\|^{\star}_{p,M} \Big) .$$

Let us now assume that $v \mapsto DC[v]$ and $(v, Y) \mapsto D\boldsymbol{F}[v, Y]$ are Lipschitz with constant L', in the sense that for all pairs (p, M) as in (5.3.3), and all $\xi \in E_1$ [26]

$$\left\|(DC[v]-DC[u])\cdot\xi\right\|^{\star}_{p,M} \le L' \cdot \|v-u\|_E \cdot \|\xi\|_E \tag{5.3.8}$$

and
$$\left\|\big(D\boldsymbol{F}[v, X[u]]-D\boldsymbol{F}\big[u, X[u]\big]\big)\cdot\xi\right\|^{\star}_{p,M} \le L' \cdot \|v-u\|_E \cdot \|\xi\|_E . \tag{5.3.9}$$

Then an application of proposition 5.2.22 on page 295 produces

$$\big\|DX[v] - DX[u]\big\| \le const \cdot \|v-u\|_E , \tag{5.3.10}$$

where $\|\ \|$ denotes the operator norm on $DX[u] : E \to \mathfrak{S}^{\star n}_{p,M}$.

Let us specialize to the situation that $E = \mathbb{R}^k$. Then, by letting ξ run through the usual basis, we see that $DX[u]$ can be identified with an $n\times k$-matrix-valued process in $\mathfrak{D}^{n\times k}$. At this juncture it is necessary to ***assume that***[27] $\boldsymbol{q} > \boldsymbol{p} > \boldsymbol{k}$. Corollary 5.2.23 then puts us in the following situation:

5.3.8 *For every $\omega \in \Omega$, $u \mapsto DX[u]_{\boldsymbol{.}}(\omega)$ is a continuous map*[21] *from U to $\mathscr{D}^{n\times k}$.*

Consider now a curve $\gamma : [0,1] \to U$ that is piecewise[28] of class[10] C^1. Then the integral

$$\int_\gamma DX[u]\, d\tau \stackrel{\text{def}}{=} \int_0^1 DX[\gamma(\tau)] \cdot \gamma'(\tau)\, d\tau$$

---

[26] E_1 is the unit ball of E.
[27] See however theorem 5.3.10 below.
[28] That is to say, there exists a càglàd function $\gamma' : [0,1] \to E$ with finitely many discontinuities so that $\gamma(t) = \int_0^t \gamma'(\tau)\, d\tau \in U \quad \forall\, t \in [0,1]$.

can be understood in two ways: as the Riemann integral[29] of the piecewise continuous curve $t \mapsto DX[\gamma(\tau)] \cdot \gamma'(\tau)$ in $\mathfrak{S}^{\star n}_{p,M}$, yielding an element of $\mathfrak{S}^{\star n}_{p,M}$ that is unique up to evanescence; or else as the Riemann integral of the piecewise continuous curve $\tau \mapsto DX[\gamma(\tau)]_{\bullet}(\omega) \cdot \gamma'(\tau)$, one for every $\omega \in \Omega$, and yielding for every $\omega \in \Omega$ an element of path space $\mathscr{D}^n$. Looking at Riemann sums that approximate the integrals will convince the reader that the integral understood in the latter sense is but one of the (many nearly equal) processes that constitute the integral of the former sense, which by the Fundamental Theorem of Calculus equals $X[\gamma(1)]_{\bullet} - X[\gamma(0)]_{\bullet}$. In particular, if γ is a closed curve, then the integral in the first sense is evanescent; and this implies that for nearly every $\omega \in \Omega$ the Riemann integral in the second sense,

$$\oint_\gamma DX[u]_{\bullet}(\omega)\, d\tau \tag{$*$}$$

is the zero path in $\mathscr{D}^n$.

Now let Γ denote the collection of all closed polygonal paths in $U \subseteq \mathbb{R}^k$ whose corners are rational points. Clearly Γ is countable. For every $\gamma \in \Gamma$ the set of $\omega \in \Omega$ where the integral $(*)$ is non-zero is nearly empty, and so is the union of these sets. Let us remove it from Ω. That puts us in the position that $\oint_\gamma DX[u](\omega)\, d\tau = 0$ for all $\gamma \in \Gamma$ and all $\omega \in \Omega$. Now for every curve that is piecewise[28] of class C^1 there is a sequence of curves $\gamma_n \in \Gamma$ such that both $\gamma_n(\tau) \to \gamma_n(\tau)$ and $\gamma'_n(\tau) \to \gamma'_n(\tau)$ uniformly in $\tau \in [0,1]$. From this it is plain that the integral $(*)$ vanishes for every closed curve γ that is piecewise of class C^1, on every $\omega \in \Omega$.

To bring all of this to fruition, let us pick in every component U_0 of U a base point u_0 and set, for every $\omega \in \Omega$,

$$X[u]_{\bullet}(\omega) \stackrel{\text{def}}{=} \int_\gamma DX[\gamma(\tau)]_{\bullet}(\omega) \cdot \gamma'(\tau)\, d\tau\ ,$$

where γ is some C^1-path joining u_0 to $u \in U_0$. This element of $\mathscr{D}^n$ does not depend on γ, and $X[u]_{\bullet}$ is one of the (many nearly equal) solutions of our stochastic differential equation (5.3.1). The upshot:

Proposition 5.3.9 *Assume that the initial condition and the coupling coefficients of equation (5.3.1) on page 298 have weak derivatives $DC[u]$ and $DF_\eta[u,X]$ in $\mathfrak{S}^{\star n}_{p,M}$ that are Lipschitz in their argument u, in the sense of (5.3.8) and (5.3.9), respectively. Assume further that $\boldsymbol{Z}$ is a local L^q-integrator for some $q > \dim U$. Then there exists a particular solution $X[u]_{\bullet}(\omega)$ that is, for nearly every $\omega \in \Omega$, differentiable as a map from U to path space[21] $\mathscr{D}^n$.*

Using theorem 5.2.24 on page 295, it suffices to assume that $\boldsymbol{Z}$ is an L^0-integrator when $\boldsymbol{F}$ is an autologous coupling coefficient:

[29] See exercise A.3.16 on page 401.

Theorem 5.3.10 *Suppose that the F_η are differentiable in the sense of example 5.3.2, their derivatives being Lipschitz in $u \in U \subseteq \mathbb{R}^k$. Then there exists a particular solution $X[u]_.(\omega)$ that is, for every $\omega \in \Omega$, differentiable as a map from U to path space $\mathscr{D}^n$.*[17]

Higher Order Derivatives

Again let $(E, \| \ \|_E)$ and $(S, \| \ \|_S)$ be seminormed spaces and let $\mathcal{U} \subseteq E$ be open and convex. To paraphrase definition A.2.49, a function $F : \mathcal{U} \to S$ is differentiable at $u \in \mathcal{U}$ if "*it can be approximated at u by an affine function strictly better than linearly.*" We can paraphrase Taylor's formula A.2.42 similarly: a function on $U \subseteq \mathbb{R}^k$ with continuous derivatives up to order l at u "*can be approximated at u by a polynomial of degree l to an order strictly better than l.*" In fact, Taylor's formula is the main merit of having higher order differentiability. It is convenient to use this behavior as the *definition* of differentiability to higher orders. It essentially agrees with the usual recursive definition (exercise 5.3.18 on page 310).

Definition 5.3.11 *Let $\| \ \|_S^\circ \le \| \ \|_S$ be a seminorm on S that satisfies $\|x\|_S = 0 \Leftrightarrow \|x\|_S^\circ = 0 \ \forall x \in S$. The map $F : \mathcal{U} \to S$ is* ***l-times $\| \ \|_S^\circ$-weakly differentiable*** *at $u \in \mathcal{U}$ if there exist continuous symmetric λ-forms*[30]

$$D^\lambda F[u] : \underbrace{E \otimes \cdots \otimes E}_{\lambda \text{ factors}} \to S \,, \qquad \lambda = 1, \ldots, l \,,$$

such that
$$F[v] = \sum_{0 \le \lambda \le l} \frac{1}{\lambda!} D^\lambda F[u]\cdot(v-u)^{\otimes\lambda} + R^l F[u;v] \,, \tag{5.3.11}$$

where
$$\left\| R^l F[u;v] \right\|_S^\circ = o\big(\|v-u\|_E^l\big) \quad \text{as} \quad v \to u \,.$$

$D^\lambda F[u]$ is the ***λ^{th} derivative*** *of F at u; and the first sum on the right in (5.3.11) is the* ***Taylor polynomial of degree l*** *of F at u, denoted $T^l F[u] : v \mapsto T^l F[u](v)$. If*

$$\sup\left\{ \frac{\|R^l F[u;v]\|_S^\circ}{\delta^l} : u,v \in U \,,\ \|v-u\|_E < \delta \right\} \xrightarrow[\delta\to 0]{} 0 \,,$$

then F is ***l-times $\| \ \|_S^\circ$-weakly uniformly differentiable****. If the target space of F is $\mathfrak{S}_{p,M}^{\star n}$, we say that F is l-times weakly (uniformly) differentiable provided it is l-times $\| \ \|_{p^\circ,M^\circ}^\star$-weakly (uniformly) differentiable in the sense above for every Picard norm*[11] *$\| \ \|_{p^\circ,M^\circ}^\star$ with $p^\circ < p$ and $M^\circ > M$.*

---

[30] A λ-form on E is a function D of λ arguments in E that is linear in each of its arguments separately. It is ***symmetric*** if it equals its ***symmetrization***, which at $(\xi_1, \cdots, \xi_\lambda) \in E^\lambda$ has the value $\frac{1}{\lambda!} \sum D\cdot\xi_{\pi_1} \otimes \cdots \otimes \xi_{\pi_\lambda}$, the sum being taken over all permutations π of $\{1, \cdots, \lambda\}$.

In (5.3.11) we write $D^\lambda F[u]\cdot\xi_1\otimes\cdots\otimes\xi_\lambda$ for the value of the form $D^\lambda F[u]$ at the argument $(\xi_1,\ldots,\xi_\lambda)$ and abbreviate this to $D^\lambda F[u]\cdot\xi^{\otimes\lambda}$ if $\xi_1=\cdots=\xi_\lambda=\xi$. For $\lambda=0$, $D^0F[u]\cdot(v-u)^{\otimes 0}$ stands as usual for the constant $F[u]\in S$. $D^\lambda F[u]\cdot\xi_1\otimes\cdots\otimes\xi_\lambda$ can be constructed from the values $D^\lambda F[u]\cdot\xi^{\otimes\lambda}$, $\xi\in E$: it is the coefficient of $\tau_1\cdots\tau_\lambda$ in $D^\lambda F[u]\cdot(\tau_1\xi_1+\cdots+\tau_\lambda\xi_\lambda)^{\otimes\lambda}/\lambda!$. To say that $D^\lambda F[u]$ is continuous means that

$$\begin{aligned}\left\|D^\lambda F[u]\right\| &\stackrel{\text{def}}{=} \sup\left\{\left\|D^\lambda F[u]\cdot\xi_1\otimes\cdots\otimes\xi_\lambda\right\|_S : \|\xi_1\|_E\le 1,\ldots,\|\xi_\lambda\|_E\le 1\right\}\\ &\le \lambda^{\lambda/2}\sup\left\{\left\|D^\lambda F[u]\cdot\xi^{\otimes\lambda}\right\|_S : \|\xi\|_E\le 1\right\}\end{aligned} \tag{5.3.12}$$

is finite (inequality (5.3.12) is left to the reader to prove). $D^\lambda F[u]$ does not depend on l; indeed, the last $l-\lambda$ terms of the sum in (5.3.11) are $o\big(\|v-u\|_E^\lambda\big)$ if measured with $\|\ \|_S^\circ$. In particular, D^1F is the weak derivative DF of definition A.2.49.

Example 5.3.12 — Trouble Consider a function $f:\mathbb{R}\to\mathbb{R}$ that has l continuous bounded derivatives, *vis.* $f(x)=\cos x$. One hopes that composition with f, which takes ϕ to $F[\phi]\stackrel{\text{def}}{=}f\circ\phi$, might define an l-times weakly[31] differentiable map from $L^p(\mathbb{P})$ to itself. Alas, it does not. Indeed, if it did, then $D^\lambda F[\phi]\cdot\psi^{\otimes\lambda}$ would have to be multiplication of the λ^{th} derivative $f^{(\lambda)}(\phi)$ with ψ^λ. For $\psi\in L^p$ this product can be expected to lie in $L^{p/\lambda}$, but not generally in L^p. However:

If $f:\mathbb{R}^n\to\mathbb{R}^n$ has continuous bounded partial derivatives of all orders $\lambda\le l$, then $F:\phi\to F[\phi]\stackrel{\text{def}}{=}f\circ\phi$ is weakly differentiable as a map from $L^p_{\mathbb{R}^n}$ to $L^{p/\lambda}_{\mathbb{R}^n}$, for $1\le\lambda\le l$, and[1]

$$D^\lambda F[\phi]\cdot\psi_1\otimes\cdots\otimes\psi_\lambda=\frac{\partial^\lambda f(\phi)}{\partial x^{\nu_\lambda}\cdots\partial x^{\nu_1}}\times\psi_1^{\nu_1}\cdots\psi_\lambda^{\nu_\lambda}\,.$$

These observations lead to a more modest notion of higher order differentiability, which, though technical and useful only for functions that take values in L^p or in $\mathfrak{S}^{\star n}_{p,M}$, has the merit of being pertinent to the problem at hand:

Definition 5.3.13 *(i) A map $F:\mathcal{U}\to\mathfrak{S}^{\star n}_{p,M}$* ***has l tiered weak derivatives*** *if for every $\lambda\le l$ it is λ-times weakly differentiable as a map from $\mathcal{U}$ to $\mathfrak{S}^{\star n}_{p/\lambda,M\lambda}$.*

(ii) A parameter-dependent coupling coefficient $F:U\times\mathfrak{S}^{\star n}_{p,M}\to\mathfrak{S}^{\star n}_{p,M}$ with l tiered weak derivatives ***has l <u>bounded</u> tiered weak derivatives*** *if*

$$\left\|D^\lambda F_\eta[u,X]\cdot\binom{\xi_1}{\Xi_1}\otimes\cdots\otimes\binom{\xi_\lambda}{\Xi_\lambda}\right\|^\star_{p/\lambda,M\lambda}\le C\prod_{1\le j\le\lambda}\Big(\|\xi_j\|_E+\left\|\Xi_j\right\|^\star_{p/i_j,Mi_j}\Big)$$

for some constant C whenever $i_1,\ldots,i_\lambda\in\mathbb{N}$ have $i_1+\cdots+i_\lambda\le\lambda$.

[31] That F is not Fréchet differentiable, not even when $l=1$, we know from example A.2.48.

Example 5.3.14 The markovian parameter-dependent coupling coefficient $(u, X) \mapsto F[u, X] \stackrel{\text{def}}{=} f(u, X)$ of example 5.3.1 on page 299 has l bounded tiered weak derivatives provided the function f has bounded continuous derivatives of all orders $\le l$. This is immediate from Taylor's formula A.2.42.

Example 5.3.15 Example 5.3.2 on page 299 has an extension as well. Assume that the map $f : U \times \mathscr{D}^n \to \mathscr{D}^n$ has l continuous bounded Fréchet derivatives. This is to mean that for every $t < \infty$ the restriction of f to $U \times \mathscr{D}^{nt}$, $\mathscr{D}^{nt}$ the Banach space of paths stopped at t and equipped with the topology of uniform convergence, is l-times continuously Fréchet differentiable, with the norm of the λ^{th} derivative being bounded in t. Then $F[u, X]_{\bullet}(\omega) \stackrel{\text{def}}{=} f[u, X_{\bullet}(\omega)]$ again defines a parameter-dependent coupling coefficient that is l-times weakly uniformly differentiable with bounded tiered derivatives.

Theorem 5.3.16 *Assume that in equation (5.3.1) on page 298 the initial value $C[u]$ has l tiered weak derivatives on U and that the coupling coefficients $F_\eta[u, X]$ have l bounded tiered weak derivatives.*

Then the solution $X[u]$ has l tiered weak derivatives on U as well, and $DX^l[u]$ is given by equation (5.3.18) below.

Proof. By theorem 5.3.6 on page 302 this is true when $l = 1$ – a good start for an induction. In order to get an idea what the derivatives $D^\lambda X[u]$ might be when $1 < \lambda \le l$, let us assume that X does in fact have l tiered weak derivatives. With

$$v = u + \tau\xi \ , \quad \|\xi\|_E = 1 \ , \quad \tau = \|v - u\|_E \ ,$$

we write this as

$$X[v] - X[u] = \sum_{1 \le \lambda \le l} \frac{\tau^\lambda}{\lambda!} D^\lambda X[u] \cdot \xi^{\otimes\lambda} + R^l X[u; v] \ , \tag{5.3.13}$$

where[5] $$\left\| R^l X[u; v] \right\|_{p^\circ/l, M^\circ l} = o(\tau^l) \qquad \text{for } p^\circ < p \ , \ M^\circ > M \ .$$

On the other hand,

$$\begin{aligned} X[v] - X[u] &= C[v] - C[u] + \Big\{ \boldsymbol{F}[v, X[v]] - \boldsymbol{F}\big[u, X[u]\big] \Big\}_{\bullet-} * \boldsymbol{Z} \\ &= \sum_{1 \le \lambda \le l} \frac{\tau^\lambda}{\lambda!} D^\lambda C[u] \cdot \xi^{\otimes\lambda} + R^l C[u; v] \\ &\quad + \sum_{1 \le \lambda \le l} \frac{1}{\lambda!} \left\{ D^\lambda \boldsymbol{F}\big[u, X[u]\big] \cdot \begin{pmatrix} v - u \\ X[v] - X[u] \end{pmatrix}^{\otimes\lambda} \right\}_{\bullet-} * \boldsymbol{Z} \\ &\qquad + R^l \boldsymbol{F}[u, X[u]; v, X[v]]_{\bullet-} * \boldsymbol{Z} \end{aligned} \tag{5.3.14}$$

$$= \sum_{1\le\lambda\le l} \frac{\tau^\lambda}{\lambda!} D^\lambda C[u]\cdot\xi^{\otimes\lambda} + R^l C[u;v] \tag{5.3.15}$$

$$+ \sum_{1\le\lambda\le l} \frac{1}{\lambda!} \Big\{ D^\lambda \boldsymbol{F}\big[u, X[u]\big]\cdot\Delta^\lambda[\tau]\Big\}_{\bullet-} {*}\boldsymbol{Z}$$

$$+ R^l \boldsymbol{F}[u, X[u]; v, X[v]]_{\bullet-}{*}\boldsymbol{Z} \ , \tag{5.3.16}$$

where by the multinomial formula[32]

$$\Delta^\lambda[\tau] \stackrel{\text{def}}{=} \begin{pmatrix} v-u \\ X[v]-X[u] \end{pmatrix}^{\otimes\lambda} = \begin{pmatrix} \tau\xi \\ \sum_{1\le\rho\le l} \frac{\tau^\rho D^\rho X[u]\cdot\xi^{\otimes\rho}}{\rho!} + R^l X[u;v] \end{pmatrix}^{\otimes\lambda}$$

$$= \sum_{\lambda_0+\cdots+\lambda_{l+1}=\lambda} \binom{\lambda}{\lambda_0 \ldots \lambda_{l+1}} \times \frac{\tau^{\lambda_0+1\lambda_1+\cdots+l\lambda_l}}{1!\cdots l!} \times$$

$$\times \begin{pmatrix}\xi\\0\end{pmatrix}^{\otimes\lambda_0} \otimes \begin{pmatrix}0\\ D^1X[u]\cdot\xi^{\otimes 1}\end{pmatrix}^{\otimes\lambda_1} \otimes\cdots\otimes \begin{pmatrix}0\\ D^lX[u]\cdot\xi^{\otimes l}\end{pmatrix}^{\otimes\lambda_l} \otimes \begin{pmatrix}0\\ R^lX[u;v]\end{pmatrix}^{\otimes\lambda_{l+1}}$$

and where

$$\big\|R^l C\big\|_{p^\circ/l,M^\circ l} = o(\tau^l) = \big\|R^l \boldsymbol{F}\big\|_{p^\circ/l,M^\circ l} \qquad \text{for } \ p^\circ{<}p \ , \ M^\circ{>}M$$

(the arguments of R^lC and $R^l\boldsymbol{F}$ are not displayed). Line (5.3.13), and lines (5.3.15)–(5.3.16) together, each are of the form "*a polynomial in τ plus terms that are $o(\tau^l)$*" when measured in the pertinent Picard norm, which here is $\|\ \|_{p^\circ/l,M^\circ l}$. Clearly,[25] then, the coefficient of τ^l in the two polynomials must be the same:[32]

$$\frac{D^lX[u]\cdot\xi^{\otimes l}}{l!} = \frac{D^lC[u]\cdot\xi^{\otimes l}}{l!} + \Bigg\{ \sum_{1\le\lambda\le l} \frac{1}{\lambda!} D^\lambda \boldsymbol{F}\big[u, X[u]\big]$$

$$\cdot \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_l=\lambda \\ \lambda_0+1\lambda_1+\cdots+l\lambda_l=l}} \binom{\lambda}{\lambda_0 \ldots \lambda_l} \times \frac{1}{1!\cdots l!} \times \tag{5.3.17}$$

$$\times \begin{pmatrix}\xi\\0\end{pmatrix}^{\otimes\lambda_0} \otimes \begin{pmatrix}0\\ D^1X[u]\cdot\xi^{\otimes 1}\end{pmatrix}^{\otimes\lambda_1} \otimes\cdots\otimes \begin{pmatrix}0\\ D^lX[u]\cdot\xi^{\otimes l}\end{pmatrix}^{\otimes\lambda_l} \Bigg\}_{\bullet-}{*}\boldsymbol{Z} \ .$$

The term $D^lX[u]\cdot\xi^{\otimes l}$ occurs precisely once on the right-hand side, namely when $\lambda_l = 1$ and then $\lambda = 1$. Therefore the previous equation can be rewritten as a stochastic differential equation for $D^lX[u]\cdot\xi^{\otimes l}$:

$$D^lX[u]\cdot\xi^{\otimes l} = \Big(D^lC[u]\cdot\xi^{\otimes l} + (\overline{C}^l[u]\cdot\xi^{\otimes l})_{\bullet-}{*}\boldsymbol{Z}\Big)$$

$$+ \Big(D_2\boldsymbol{F}\big[u, X[u]\big]\cdot D^lX[u]\cdot\xi^{\otimes l}\Big)_{\bullet-}{*}\boldsymbol{Z} \ , \tag{5.3.18}$$

[32] Use $\binom{\xi}{D} = \binom{\xi}{0} + \binom{0}{D}$. It is understood that a term of the form $(\cdots)^{\otimes 0}$ is to be omitted.

where $D_2\boldsymbol{F}$ is the partial derivative in the X-direction (see A.2.50) and

$$\overline{C}^l[u]\cdot\xi^{\otimes l} \stackrel{\text{def}}{=} \sum_{1\le\lambda\le l} \frac{D^\lambda \boldsymbol{F}\big[u, X[u]\big]}{\lambda!} \cdot \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_{l-1}=\lambda \\ \lambda_0+1\lambda_1+\cdots+(l-1)\lambda_{l-1}=l}} \binom{\lambda}{\lambda_0 \ldots \lambda_{l-1}} \times \frac{1}{1!\cdots(l-1)!} \times$$

$$\times \binom{\xi}{0}^{\otimes\lambda_0} \otimes \binom{0}{D^1X[u]\cdot\xi^{\otimes 1}}^{\otimes\lambda_1} \otimes\cdots\otimes \binom{0}{D^{l-1}X[u]\cdot\xi^{\otimes(l-1)}}^{\otimes\lambda_{l-1}} .$$

Now by induction hypothesis, $D^iX\cdot\xi^{\otimes i}$ stays bounded in $\mathfrak{S}^{\star n}_{p/i,Mi}$ as ξ ranges over the unit ball of E and $1 \le i \le l-1$. Therefore $D^\lambda\boldsymbol{F}\big[u, X[u]\big]$ applied to any of the summands stays bounded in $\mathfrak{S}^{\star n}_{p/l,Ml}$, and then so does $\overline{C}^l[u]\cdot\xi^{\otimes l}$. Since the coupling coefficient of (5.3.18) has Lipschitz constant L, we conclude that $D^lX[u]\cdot\xi^{\otimes l}$, *defined by* (5.3.18), stays bounded in $\mathfrak{S}^{\star n}_{p/l,Ml}$ as ξ ranges over E_1.

There is a little problem here, in that (5.3.18) defines $D^lX[u]$ as an l-homogeneous map on E, but not immediately as a l-linear map on $\bigotimes_l E$. To overcome this observe that $\overline{C}[u]\cdot\xi^{\otimes l}$ is in an obvious fashion the value at $\xi^{\otimes l}$ of an l-linear map

$$\vec{\xi}^{\otimes l} \stackrel{\text{def}}{=} \xi_1\otimes\cdots\otimes\xi_l \;\mapsto\; \overline{C}[u]\cdot\vec{\xi}^{\otimes l} .$$

Replacing every $\xi^{\otimes l}$ in (5.3.18) by $\vec{\xi}^{\otimes l}$ produces a stochastic differential equation for an n-vector $D^lX[u]\cdot\vec{\xi}^{\otimes l} \in \mathfrak{S}^{\star n}_{p/l,Ml}$, whose solution defines an l-linear form that at $\vec{\xi}^{\otimes l} = \xi^{\otimes l}$ agrees with the $D^lX[u]$ of equation (5.3.18). The l^{th} derivative $D^lX[u]$ is redefined as the symmetrization[30] of this l-form. It clearly satisfies equation (5.3.18) and is the only symmetric l-linear map that does.

It is left to be shown that for $l>1$ the difference $R^l[u;v]$ of $X[v]-X[u]$ and the Taylor polynomial $T^lX[u](\tau\xi)$ is $o(\tau^l)$ if measured in $\mathfrak{S}^{\star n}_{p^\circ/l,M^\circ l}$. Now by induction hypothesis, $R^{l-1}[u;v] = D^lX[u](v-u)^{\otimes l}/l! + R^l[u;v]$ is $o(\tau^{l-1})$; hence clearly so is $R^l[u;v]$. Subtracting the defining equations (5.3.17) for $l=1,2,\ldots$ from (5.3.13) and (5.3.15)–(5.3.16) leaves us with this equation for the remainder $R^lX[u;v]$:

$$\begin{aligned} R^lX[u;v] = R^lC[u;v] &+ \sum_{1\le\lambda\le l} \frac{1}{\lambda!}\Big\{D^\lambda\boldsymbol{F}\big[u,X[u]\big]\cdot\overline{\Delta}^\lambda[\tau]\Big\}_{\cdot-}*\boldsymbol{Z} \\ &+ R^l\boldsymbol{F}[u,X[u];v,X[v]]_{\cdot-}*\boldsymbol{Z} , \end{aligned} \tag{5.3.19}$$

where

$$\overline{\Delta}^\lambda[\tau] \stackrel{\text{def}}{=} \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_l=\lambda \\ \lambda_0+1\lambda_1+\cdots+l\lambda_l>l}} \binom{\lambda}{\lambda_0\ldots\lambda_l} \times \frac{\tau^{\lambda_0+1\lambda_1+\cdots+l\lambda_l}}{1!\cdots l!} \times$$

$$\times \binom{\xi}{0}^{\otimes\lambda_0} \otimes \binom{0}{D^1X[u]{\cdot}\xi^{\otimes 1}}^{\otimes\lambda_1} \otimes\cdots\otimes \binom{0}{D^lX[u]{\cdot}\xi^{\otimes l}}^{\otimes\lambda_l}$$

$$+ \sum_{\substack{\lambda_0+\lambda_1+\cdots+\lambda_{l+1}=\lambda \\ \lambda_{l+1}>0}} \binom{\lambda}{\lambda_0\ldots\lambda_{l+1}} \times \frac{\tau^{\lambda_0+1\lambda_1+\cdots+l\lambda_l}}{1!\cdots l!} \times$$

$$\times \binom{\xi}{0}^{\otimes\lambda_0} \otimes \binom{0}{D^1X[u]{\cdot}\xi^{\otimes 1}}^{\otimes\lambda_1} \otimes\cdots\otimes \binom{0}{D^lX[u]{\cdot}\xi^{\otimes l}}^{\otimes\lambda_l} \otimes \binom{0}{R^lX[u;v]}^{\otimes\lambda_{l+1}} .$$

The terms in the first sum all are $o(\tau^l)$. So are all of the terms of the second sum, except the one that arises when $\lambda_{l+1}=1$ and $\lambda_0+1\lambda_1+\cdots+l\lambda_l=0$ and then $\lambda=1$. That term is $\binom{0}{R^lX[u;v]/l!}$. Lastly, $R^l\boldsymbol{F}[u,X[u];v,X[v]]$ is easily seen to be $o(\tau^l)$ as well. Therefore equation (5.3.19) boils down to a stochastic differential equation for $R^lX[u;v]$:

$$R^lX[u;v] = \Big\{R^lC[u;v] + o(\tau^l)_{.-}{*}\boldsymbol{Z}\Big\} + \big(D_2F\big[u,X[u]\big]{\cdot}R^lX[u;v]\big)_{.-}{*}\boldsymbol{Z} .$$

According to inequalities (5.2.23) on page 290 and (5.2.5) on page 284, we have $R^lX[u;v] = o(\|u-v\|_E^l)$, as desired. ∎

Exercise 5.3.17 If in addition C and $\boldsymbol{F}$ are weakly uniformly differentiable, then so is X.

Exercise 5.3.18 Suppose $F : U \to S$ is l-times weakly uniformly differentiable with bounded derivatives:

$$\sup_{\lambda\le l\,,\,u\in U} \Big\| D^\lambda F[u] \Big\| < \infty \qquad \text{(see inequality (5.3.12)).}$$

Then, for $\lambda < l$, $D^\lambda F$ is weakly uniformly differentiable, and its derivative is $D^{\lambda+1}F$.

Problem 5.3.19 Generalize the pathwise differentiability result 5.3.9 to higher order derivatives.

5.4 Pathwise Computation of the Solution

We return to the stochastic differential equation (5.1.3), driven by a vector $\boldsymbol{Z}$ of integrators:

$$X = C + F_\eta[X]_{.-}{*}Z^\eta = C + \boldsymbol{F}[X]_{.-}{*}\boldsymbol{Z} .$$

Under mild conditions on the coupling coefficients F_η there exists an algorithm that computes the path $X_.(\omega)$ of the solution from the input paths $C_.(\omega), \boldsymbol{Z}_.(\omega)$. It is a slight variant of the well-known adaptive[33] Euler–Peano

[33] Adaptive: the step size is not fixed in advance but is adapted to the situation at every step – see page 281.

scheme of little straight steps, a variant in which the next computation is carried out not after a fixed time has elapsed but when the effect of the noise $\boldsymbol{Z}$ has changed by a fixed threshold – compare this with exercise 3.7.24. There exists an algorithm that takes the $n+d$ paths $t \mapsto C_t^\nu(\omega)$ and $t \mapsto Z_t^\eta(\omega)$ and computes from them a path $t \mapsto {}^\delta X_t(\omega)$, which, when δ is taken through a summable sequence, converges ω–by–ω uniformly on bounded time-intervals to the path $t \mapsto X_t(\omega)$ of the exact solution, irrespective of $\mathbb{P} \in \mathfrak{P}[\boldsymbol{Z}]$. This is shown in theorems 5.4.2 and 5.4.5 below.

The Case of Markovian Coupling Coefficients

One cannot of course expect such an algorithm to exist unless the coupling coefficients F_η are endogenous. This is certainly guaranteed when the coupling coefficients are markovian[8], case treated first. That is to say, we assume here that there are ordinary vector fields $f_\eta : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$F_\eta[X]_t = f_\eta \circ X_t \quad ; \text{and} \quad |f_\eta(y) - f_\eta(x)| \le L \cdot |y - x| \tag{5.4.1}$$

ensures the Lipschitz condition (5.2.8) and with it the existence of a unique solution of equation (5.2.1), which takes the following form:[1]

$$X_t = C_t + \int_0^t f_\eta(X)_{s-} \, dZ_s^\eta \ . \tag{5.4.2}$$

The adaptive[33] Euler–Peano algorithm computing the approximate X' for a fixed threshold[34] $\delta > 0$ works as follows: define $T_0 \stackrel{\text{def}}{=} 0$, $X'_0 \stackrel{\text{def}}{=} C_0$ and continue recursively: when the stopping times $T_0 \le T_1 \le \ldots \le T_k$ and the function $X' : [\![0, T_k]\!] \to \mathbb{R}$ have been defined so that $X'_{T_k} \in \mathcal{F}_{T_k}$, then set[1]

$${}^0\Xi'_t \stackrel{\text{def}}{=} C_t - C_{T_k} + f_\eta\big(X'_{T_k}\big) \cdot \big(Z_t^\eta - Z_{T_k}^\eta\big) \tag{5.4.3}$$

and
$$T_{k+1} \stackrel{\text{def}}{=} \inf\Big\{ t > T_k : \big|{}^0\Xi'_t\big| > \delta \Big\} \ , \tag{5.4.4}$$

and extend X':
$$X'_t \stackrel{\text{def}}{=} X'_{T_k} + {}^0\Xi'_t \qquad \text{for } T_k \le t \le T_{k+1} \ . \tag{5.4.5}$$

In other words, the prescription is to wait after time T_k not until some fixed time has elapsed but until random input plus effect of the drivers together have changed sufficiently to warrant a new computation; then extend X' "linearly" to the interval that just passed, and start over. It is obvious how to write a little loop for a computer that will compute the path $X'_{\boldsymbol{.}}(\omega)$ of the ***Euler–Peano approximate*** $X'_{\boldsymbol{.}}(\omega)$ as it receives the input paths $C_{\boldsymbol{.}}(\omega)$ and $\boldsymbol{Z}_{\boldsymbol{.}}(\omega)$. The scheme (5.4.5) expresses quite intuitively the meaning of the differential equation $dX = \boldsymbol{f}(X)\, d\boldsymbol{Z}$. If one can show that it converges, one should be satisfied that the limit is for all intents and purposes a solution of the differential equation (5.4.2).

---

[34] Visualize δ as a ***step size*** on the dependent variables' axis.

One can, and it is. An easy induction shows that[1,35]

$$\text{for} \qquad t < T_\infty \stackrel{\text{def}}{=} \sup_{k<\infty} T_k$$

$$\text{we have} \qquad X'_t = C_t + \sum_{k=0}^{\infty} f_\eta\big(X'_{T_k}\big) \cdot \big(Z^\eta_{T_{k+1}\wedge t} - Z^\eta_{T_k\wedge t}\big)$$

$$\text{by 3.5.2:} \qquad = C_t + \int_0^t \sum_{0\le k<\infty} f_\eta\big(X'_{T_k}\big) \cdot (\!(T_k, T_{k+1}]\!]\, dZ^\eta \tag{5.4.6}$$

and exhibits $X' : [\![0, T_\infty)\!) \to \mathbb{R}$ as an adapted process that is right-continuous with left limits on $[\![0, T_\infty)\!)$ – but for all we know so far not necessarily at T_∞.

On the way to proving that the Euler–Peano approximate X' is close to the exact solution X, the first order of business is to show that $T_\infty = \infty$. In order to do this recall the scalæfication of processes in definition 3.7.22 that is attached to the random partition[36] $\mathcal{T} \stackrel{\text{def}}{=} \{0 = T_0 \le T_1 \le \ldots \le T_\infty \le \infty\}$:

$$\text{for } Y \in \mathfrak{D}^n, \qquad Y^{\mathcal{T}} \stackrel{\text{def}}{=} \sum_{0\le k\le\infty} Y_{T_k} \cdot [\![T_k, T_{k+1})\!) \;\; \in \mathfrak{D}^n \,.$$

Attach now a new coupling coefficient $\boldsymbol{F}'$ to $\boldsymbol{F}$ and the partition $\mathcal{T}$ by

$$F'_\eta[Y] \stackrel{\text{def}}{=} F_\eta[Y^{\mathcal{T}}]^{\mathcal{T}} \qquad \text{for } Y \in \mathfrak{D}^n \text{ and } \eta = 1, \ldots, d, \tag{5.4.7}$$

and consider the stochastic differential equation[1]

$$Y = C + F'_\eta[Y]_{.-} {*} Z^\eta \,, \tag{5.4.8}$$

$$\text{which reads}^{35} \qquad Y_t = C_t + \int_0^t \sum_{0\le k\le\infty} f_\eta\big(Y_{T_k}\big) \cdot (\!(T_k, T_{k+1}]\!]\, dZ^\eta \tag{5.4.9}$$

in the present markovian case. The coupling coefficient $\boldsymbol{F}'$ evidently satisfies the Lipschitz condition (5.2.8) with Lipschitz constant L from (5.4.1). There is therefore a *unique global* solution Y to equation (5.4.8), and a comparison of (5.4.9) with equation (5.4.6) reveals that $X' = Y$ on $[\![0, T_\infty)\!)$. On the set $[T_\infty < \infty]$, $Y \in \mathfrak{D}^n$ has almost surely no oscillatory discontinuity, but X' surely does, since the values of this process at T_k and T_{k+1} differ by at least δ, yet $\sup_k |X'|_{T_k}$ is bounded by $|Y|^\star_{T_\infty} < \infty$. The set $[T_\infty < \infty]$ is therefore negligible, even nearly empty, and thus the T_k increase without bound, nearly.

---

35 In accordance with convention A.1.5 on page 364, sets are identified with their (idempotent) indicator functions. A stochastic interval $(\!(S, T]\!]$, for instance, has at the instant s the value $(\!(S, T]\!]_s = [S < s \le T] = \begin{cases} 1 & \text{if } S(\omega) < s \le T(\omega) \\ 0 & \text{elsewhere} \end{cases}$.

36 A partition $\mathcal{T}$ is assumed to contain $T_\infty \stackrel{\text{def}}{=} \sup_{k<\infty} T_k$, and the convention $T_{\infty+1} \stackrel{\text{def}}{=} \infty$ simplifies formulas.

We now run Picard's iterative scheme, starting with the scalæfication[35]

$$X^{(0)} \stackrel{\text{def}}{=} X'^{\mathcal{T}} = \sum_{k=0}^{\infty} X'_{T_k} \cdot [\![T_k, T_{k+1})\!) \ .$$

Then
$$X^{(1)} \stackrel{\text{def}}{=} \mathfrak{U}[X^{(0)}] = C + f_\eta(X^{(0)})_{.-} {*} Z^\eta$$

in view of (5.4.6) equals X' and differs from $X^{(0)}$ by less than δ uniformly on $\boldsymbol{B}$. Therefore $X^{(0)}$ and $X^{(1)}$ differ by less than δ in any of the norms $\| \ \|^\star_{p,M}$. The argument of item 5.1.5 immediately provides the estimate (i) below. (Another way to arrive at inequality (5.4.10) is to observe that, on the solution X' of (5.4.8), the F_η and F'_η differ by less than $\delta{\cdot}L$, and to invoke inequality (5.2.35).)

Proposition 5.4.1 *Assume that $\boldsymbol{Z}$ is a local L^q-integrator for some $q \ge 2$ and that equation (5.4.2) with markovian coupling coefficients as in (5.4.1) has its exact global solution X inside $\mathfrak{S}^{\star n}_{p,M}$ for some[23] $p \in [2, q]$ and some[23] $M > M^{\diamond(5.2.20)}_{p,L}$. Then, with $\gamma \stackrel{\text{def}}{=} M^{\diamond(5.2.20)}_{p,L}/M$,*

$$\text{(i)} \qquad \big\|X - X'\big\|^\star_{p,M} \le \frac{\gamma}{1-\gamma} \cdot \delta\ ; \tag{5.4.10}$$

$$\text{(ii) and} \qquad \big|X - X^{(n)}\big|^\star_T \xrightarrow[n\to\infty]{} 0 \quad \text{almost surely}$$

at any nearly finite stopping time T, whenever the $X^{(n)}$ are the Euler–Peano approximates constructed via (5.4.5) from a summable sequence of thresholds $\delta_n > 0$. In other words, for nearly every $\omega \in \Omega$ we have $X^{(n)}_t \xrightarrow[n\to\infty]{} X_t$, uniformly in $t \in [0, T(\omega)]$.

Statement (ii) does not change if $\mathbb{P}$ is replaced with an equivalent probability $\mathbb{P}'$ in the manner of the proof of theorem 5.2.15; thus it holds without assuming more about $\boldsymbol{Z}$ than that it be an L^0-integrator:

Theorem 5.4.2 *Let X denote the strong global solution of the markovian system (5.4.1), (5.4.2), driven by the L^0-integrator $\boldsymbol{Z}$. Fix any summable sequence of strictly positive reals δ_n and let $X^{(n)}$ be defined as the Euler–Peano approximates of (5.4.5) for $\delta = \delta_n$. Then $X^{(n)} \xrightarrow[n\to\infty]{} X$ uniformly on bounded time-intervals, nearly.* ∎

Proof of Proposition 5.4.1 (ii). This is a standard application of the Borel–Cantelli lemma. Namely, suppose that T is one of the T^λ, say $T = T^\mu$. Then

$$\big\|X - X^{(n)}\big\|^\star_{p,M} \le \delta_n \cdot \frac{\gamma}{1-\gamma}$$

implies
$$\Big\| |X - X^{(n)}|^\star_{T^\mu-} \Big\|_{L^p} \le \delta_n \cdot \frac{e^{M\mu}\gamma}{1-\gamma}\ ,$$

and
$$\mathbb{P}\Big[|X - X^{(n)}|^\star_{T^\mu-} > \sqrt{\delta_n}\Big] \le \Big(\frac{\delta_n \times e^{M\mu}\gamma}{\sqrt{\delta_n}(1-\gamma)}\Big)^p = const \cdot \delta_n^{p/2}\ ,$$

which is summable over n, since $p \ge 2$. Therefore

$$\mathbb{P}\Big[\limsup_n |X - X^{(n)}|^\star_{T^\mu -} > 0\Big] = 0 \,.$$

For arbitrary almost surely finite T, the set $[\limsup_n |X - X^{(n)}|^\star_t > 0]$ is therefore almost surely a subset of $[T \ge T^\mu]$ and is negligible since T^μ can be made arbitrarily large by the choice of μ.

Remark 5.4.3 In the adaptive[33] Euler–Peano algorithm (5.4.5) on page 311, any stochastic partition $\mathcal{T}$ can replace the specific partition (5.4.4), as long as[36] $T_\infty = \infty$ and the quantity ${}^0\Xi'_t$ does not change by more than δ over its intervals. Suppose for instance that C is constant and the f_η are bounded, say[4] $|f_\eta(x)| \le K$. Then the partition defined recursively by $0 = T_0$, $T_{k+1} \stackrel{\text{def}}{=} \inf\{t > T_k : \sum_\eta |Z^\eta_t - Z^\eta_{T_k}| > \delta/K\}$ will do.

The Case of Endogenous Coupling Coefficients

For any algorithm similar to (5.4.5) and intended to apply in more general situations than the markovian one treated above, the coupling coefficients F_η still must be special. Namely, given *any* input path $(x_., \boldsymbol{z}_.)$, F_η must return an output path. That is to say, the F_η must be endogenous Lipschitz coefficients as in example 5.2.12 on page 289. If they are, then in terms of the f_η the system (5.1.5) reads

$$X_t = C_t + \sum_\eta \int_0^t f_\eta[\boldsymbol{Z}_., X_.]_{s-}\, dZ^\eta_s \tag{5.4.11}$$

or, equivalently, $X = C + \sum_\eta f_\eta[\boldsymbol{Z}_., X_.]_{.-} {*} Z^\eta = C + \boldsymbol{f}[\boldsymbol{Z}_., X_.]_{.-} {*} \boldsymbol{Z}$. (5.4.12)

The adaptive[33] Euler–Peano algorithm (5.4.5) needs to be changed a little. Again we fix a strictly positive threshold δ, set $T_0 \stackrel{\text{def}}{=} 0$, $X'_0 \stackrel{\text{def}}{=} C_0$, and continue recursively: when the stopping times $T_0 \le T_1 \le \ldots \le T_k$ and the function $X' : [\![0, T_k]\!] \to \mathbb{R}$ have been defined so that $X'_{T_k} \in \mathcal{F}_{T_k}$, then set[1,3]

$${}^0\boldsymbol{f}_t \stackrel{\text{def}}{=} \sup_{\eta,\nu} \big| f^\nu_\eta[\boldsymbol{Z}^{T_k}, X'^{T_k}]_t - f^\nu_\eta[\boldsymbol{Z}^{T_k}, X'^{T_k}]_{T_k} \big| \,, \quad t \ge T_k \,;$$

$${}^0\Xi'_t \stackrel{\text{def}}{=} C_t - C_{T_k} + f_\eta\big[\boldsymbol{Z}^{T_k}, X'^{T_k}\big]_{T_k} \cdot \big(Z^\eta_t - Z^\eta_{T_k}\big) \,, t \ge T_k \,;$$

and $$T_{k+1} \stackrel{\text{def}}{=} \inf\{t > T_k : {}^0\boldsymbol{f}_t > \delta \ \text{ or } \ |{}^0\Xi'_t| > \delta\} \,;$$

and then extend $X' X'_t \stackrel{\text{def}}{=} X'_{T_k} + {}^0\Xi'_t$ for $T_k \le t \le T_{k+1}$. (5.4.13)

The spirit is that of (5.4.5), the stopping times T_k are possibly "a bit closer together than there," to make sure that $\boldsymbol{f}[\boldsymbol{Z}, X']_.$ does not vary too much on the intervals of the partition[36] $\mathcal{T} \overset{\text{def}}{=} \{T_0 \le T_1 \le \ldots \le \infty\}$. An induction shows as before that

$$\text{for} \qquad t < T_\infty \overset{\text{def}}{=} \sup_{k<\infty} T_k$$

$$\text{we have} \qquad X'_t = C_t + \sum_{k=0}^{\infty} f_\eta\big[\boldsymbol{Z}, X'\big]_{T_k} \cdot \big(Z^\eta_{T_{k+1}\wedge t} - Z^\eta_{T_k\wedge t}\big)$$

$$= C_t + \int_0^t F'_\eta[\boldsymbol{Z}, X']_{\centerdot-} dZ^\eta ,$$

where the "$\mathcal{T}$-scalæfied" coupling coefficient F'_η is defined as in (5.4.7), for the present partition $\mathcal{T}$ of course. This exhibits $X' : [\![0, T_\infty)\!) \to \mathbb{R}$ as an adapted process that is right-continuous with left limits on $[\![0, T_\infty)\!)$ – but for all we know so far not necessarily at T_∞. Again we consider the stochastic differential equation (5.4.8):

$$Y = C + F'_\eta[Y]_{\centerdot-} {*} Z^\eta$$

and see that X' agrees with its unique global solution Y on $[\![0, T_\infty)\!)$. As in the markovian case we conclude from this that X' has no oscillatory discontinuities. Clearly $f_\eta[\boldsymbol{Z}^{\mathcal{T}}, X'^{\mathcal{T}}]$, which agrees with $f_\eta[\boldsymbol{Z}^{\mathcal{T}}, Y^{\mathcal{T}}]$ on $[\![0, T_\infty)\!)$, has no oscillatory discontinuities either. On the other hand, the very definition of T_∞ implies that one or the other of these processes surely must have a discontinuity on $[T_\infty < \infty]$. This set therefore is negligible, and $T_\infty = \infty$ almost surely.

Let us define $X^{(0)} \overset{\text{def}}{=} X'^{\mathcal{T}}$ and $X^{(1)} \overset{\text{def}}{=} \mathfrak{U}\big[X^{(0)}\big]$. Then

$$X^{(1)}_t = C_t + \int_0^t f_\eta\big[\boldsymbol{Z}, X^{(0)}\big]_{\centerdot-} \, dZ^\eta$$

$$\text{and} \qquad X'_t = C_t + \int_0^t f_\eta\big[\boldsymbol{Z}, X^{(0)}\big]^{\mathcal{T}}_{\centerdot-} \, dZ^\eta$$

differ by less than $\delta C^\diamond_p/eM$ when measured with the norm $\|\ \|^\star_{p,M}$ (see exercise 5.2.2). X' and $X^{(0)}$ differ uniformly, and therefore in $\|\ \|^\star_{p,M}$, by less than δ. Therefore

$$\big\|X^{(1)} - X^{(0)}\big\|^\star_{p,M} \le \delta + \delta C^\diamond_p/eM ,$$

and in view of item 5.1.4 on page 276 the approximate $X^{(1)}$ differs little from the exact solution X of (5.2.1); in fact,

$$\big\|X - X^{(1)}\big\|^\star_{p,M} \le \delta \cdot \frac{M + C^\diamond_p/e}{M - M^\diamond} ,$$

$$\text{and so} \qquad \big\|X - X'\big\|^\star_{p,M} \le \delta \cdot \frac{M + 2C^\diamond_p/e}{M - M^\diamond} . \tag{5.4.14}$$

We have recovered proposition 5.4.1 in the present non-markovian setting:

Proposition 5.4.4 *Assume that* $\boldsymbol{Z}$ *is a local* L^q*-integrator for some* $q \geq 2$, *and pick a* $p \in [2, q]$ *and an* $M > M^{\diamond(5.2.20)}_{p,L}$, L *being the Lipschitz constant of the endogenous coefficient* $\boldsymbol{f}$. *Then the global solution* X *of the Lipschitz system (5.4.11) lies in* $\mathfrak{S}^{\star n}_{p,M}$, *and the Euler–Peano approximate* X' *defined in equation (5.4.13) satisfies inequality (5.4.14).*

Even if $\boldsymbol{Z}$ is merely an L^0-integrator, this implies as in theorem 5.4.2 *the*

Theorem 5.4.5 *Fix any summable sequence of strictly positive reals* δ_n *and let* $X^{(n)}$ *be the Euler–Peano approximates of (5.4.13) for* $\delta = \delta_n$. *Then at any almost surely finite stopping time* T *and for almost all* $\omega \in \Omega$ *the sequence* $X^{(n)}_t(\omega)$ *converges to the exact solution* $X_t(\omega)$ *of the Lipschitz system (5.4.11) with endogenous coefficients, uniformly for* $t \in [0, T(\omega)]$.

Corollary 5.4.6 *Let* $\boldsymbol{Z}, \boldsymbol{Z}'$ *be* L^0*-integrators and* X, X' *solutions of the Lipschitz systems*

$$X = C + \boldsymbol{f}[\boldsymbol{Z}_{\bullet}, X_{\bullet}]_{\cdot-} * \boldsymbol{Z} \quad \textit{and} \quad X' = C' + \boldsymbol{f}[\boldsymbol{Z}_{\bullet}, X'_{\bullet}]_{\cdot-} * \boldsymbol{Z}'$$

with endogenous coefficients, respectively. Let Ω_0 *be a subset of* Ω *and* $T : \Omega \to \mathbb{R}_+$ *a time, neither of them necessarily measurable. If* $C = C'$ *and* $\boldsymbol{Z} = \boldsymbol{Z}'$ *up to and including (excluding) time* T *on* Ω_0, *then* $X = X'$ *up to and including (excluding) time* T *on* Ω_0, *except possibly on an evanescent set.*

The Universal Solution

Consider again the endogenous system (5.4.12), reproduced here as

$$X = C + \boldsymbol{f}[\boldsymbol{Z}_{\bullet}, X_{\bullet}]_{\cdot-} * \boldsymbol{Z} \, . \tag{5.4.15}$$

In view of Items 2.3.8–2.3.11, the solution can be computed on canonical path space. Here is how. Identify the process $R_t \stackrel{\text{def}}{=} (C_t, \boldsymbol{Z}_t) : \Omega \to \mathbb{R}^{n+d}$ with a representation $\underline{R}$ of $(\Omega, \mathcal{F}_{\bullet})$ on the canonical path space $\overline{\Omega} \stackrel{\text{def}}{=} \mathscr{D}^{n+d}$ equipped with its natural filtration $\overline{\mathcal{F}}_{\bullet} \stackrel{\text{def}}{=} \mathcal{F}_{\bullet}[\mathscr{D}^{n+d}]$. For consistency's sake let us denote the evaluation processes on $\overline{\Omega}$ by $\overline{\boldsymbol{Z}}$ and $\overline{C}$; to be precise, $\overline{\boldsymbol{Z}}_t(c_{\bullet}, \boldsymbol{z}_{\bullet}) \stackrel{\text{def}}{=} \boldsymbol{z}_t$ and $\overline{C}_t(c_{\bullet}, \boldsymbol{z}_{\bullet}) \stackrel{\text{def}}{=} c_t$. We contemplate the stochastic differential equation

$$\overline{X} = \overline{C} + \boldsymbol{f}[\overline{\boldsymbol{Z}}_{\bullet}, \overline{X}_{\bullet}]_{\cdot-} * \overline{\boldsymbol{Z}} \tag{5.4.16}$$

or – see (2.3.11) $\qquad \overline{X}_t(c_{\bullet}, \boldsymbol{z}_{\bullet}) = c_t + \displaystyle\int_0^t f_\eta[\boldsymbol{z}_{\bullet}, \overline{X}_{\bullet}]_{s-} \, dz^\eta_s \, .$

We produce a particularly pleasant solution $\overline{X}$ of equation (5.4.16) by applying the Euler–Peano scheme (5.4.13) to it, with $\delta = 2^{-n}$. The corresponding Euler–Peano approximate $\overline{X}^n$ in (5.4.13) is clearly adapted to the natural filtration $\mathcal{F}_{\bullet}[\mathscr{D}^{n+d}]$ on path space. Next we set $\overline{X} \stackrel{\text{def}}{=} \lim \overline{X}^n$ where this limit exists and $\overline{X} \stackrel{\text{def}}{=} 0$ elsewhere. Note that no probability enters the definition of $\overline{X}$. Yet the process $\overline{X}$ we arrive at solves the stochastic differential equation (5.4.16) in the sense of any of the probabilities in $\mathfrak{P}[\overline{\boldsymbol{Z}}]$. According to equation (2.3.12), $X_t \stackrel{\text{def}}{=} \overline{X}_t \circ \underline{R} = \overline{X}_t(C_{\bullet}, \boldsymbol{Z}_{\bullet})$ solves (5.4.15) in the sense of any of the probabilities in $\mathfrak{P}[\boldsymbol{Z}]$.

Summary 5.4.7 *The process $\overline{X}$ is càdlàg and adapted to $\mathcal{F}[\mathscr{D}^{n+d}]$, and it solves (5.4.16). Considered as a map from $\mathscr{D}^{n+d}$ to $\mathscr{D}^n$, it is adapted to the filtrations $\mathcal{F}_{\bullet}[\mathscr{D}^{n+d}]$ and $\mathcal{F}^0_{\bullet+}[\mathscr{D}^n]$ on these spaces. Since the solution X of (5.4.15) is given by $X_t = \overline{X}_t(C_{\bullet}, \boldsymbol{Z}_{\bullet})$, no matter which of the $\mathbb{P} \in \mathfrak{P}[\boldsymbol{Z}]$ prevails at the moment, $\overline{X}$ deserves the name* ***universal solution****.* ▟

A Non-Adaptive Scheme

It is natural to ask whether perhaps the stopping times T_k in the Euler–Peano scheme on page 311 can be chosen in advance, without the employ of an "infimum–detector" as in definition (5.4.4). In other words, we ask whether there is a non-adaptive scheme[33] doing the same job.

Consider again the markovian differential equation (5.4.2) on page 311:

$$X_t = C_t + \int_0^t f_\eta(X_{s-})\, dZ_s^\eta \tag{5.4.17}$$

for a vector $X \in \mathbb{R}^n$. We assume here without loss of generality that $\boldsymbol{f}(0) = 0$, replacing C with ${}^0C \stackrel{\text{def}}{=} C + \boldsymbol{f}(0)*\boldsymbol{Z}$ if necessary (see page 272). This has the effect that the Lipschitz condition (5.2.13):

$$|f_\eta(y) - f_\eta(x)| \le L \cdot |y - x| \qquad \text{implies} \qquad |f(x)| \le L \cdot |x| \,. \tag{5.4.18}$$

To simplify life a little, let us also assume that $\boldsymbol{Z}$ is quasi-left-continuous. Then the intrinsic time Λ, and with it THE time transformation $T^{\bullet}$, can and will be chosen strictly increasing and continuous (see remark 4.5.2).

Remark 5.4.8 Let us see what can be said if we simply define the T_k as usual in calculus by $T_k \stackrel{\text{def}}{=} k\delta$, $k = 0, 1, \ldots$, $\delta > 0$ being the step size. Let us denote by $\mathcal{T} = \mathcal{T}^{(\delta)}$ the (sure) partition so obtained. Then the Euler–Peano approximate X' of (5.4.5) or (5.4.6), defined by $X'_0 \stackrel{\text{def}}{=} C_0$ and recursively for $t \in (\!(T_k, T_{k+1}]\!]$ by[1]

$$X'_t \stackrel{\text{def}}{=} X'_{T_k} + \big(C_t - C_{T_k}\big) + f_\eta\big(X'_{T_k}\big) \cdot \big(Z_t^\eta - Z_{T_k}^\eta\big) \,, \tag{5.4.19}$$

is again the solution of the stochastic differential equation (5.4.8). Namely,

$$X' = C + F'_\eta[X']_{\bullet-} * Z^\eta \,,$$

with F'_η as in (5.4.7), to wit,

$$F'_\eta[Y] \stackrel{\text{def}}{=} F_\eta[Y^{\mathcal{T}}]^{\mathcal{T}} = \sum_{0 \le k < \infty} f_\eta(Y_{T_k}) \cdot [\![T_k, T_{k+1})\!) \,.$$

The stability estimate (5.2.34) on page 294 says that

$$\left\| X' - X \right\|^{\star}_{p,M} \leq \frac{\gamma}{L(1-\gamma)} \cdot \left\| \left| \boldsymbol{F}'[X] - \boldsymbol{F}[X] \right|_{\infty} \right\|_{p,M} \tag{5.4.20}$$

for any choice of $\gamma \in (0,1)$ and any $M > M^{(5.2.26)}_{L:\gamma}$. A straightforward application of the Dominated Convergence Theorem to the right-hand side of (5.4.20) shows that $\left\| X' - X \right\|^{\star}_{p,M} \to 0$ as $\delta \to 0$. Thus X' is an approximate solution, and the path of X', which is an *algebraic construct* of the path of $(C, \boldsymbol{Z})$, converges uniformly on bounded time-intervals to the path of the exact solution X as $\delta \to 0$, *in probability.*

However, the Dominated Convergence Theorem provides no control of the speed of the convergence, and this line of argument cannot rule out the possibility that *convergence* $X'_{\bullet}(\omega) \xrightarrow[\delta\to 0]{} X_{\bullet}(\omega)$ *may obtain for no single course-of-history* $\omega \in \Omega$. True, by exercise A.8.1 (iv) there exists some sequence (δ_n) along which convergence occurs almost surely, but it cannot generally be specified in advance.

A small refinement of the argument in remark 5.4.8 does however result in the desired approximation scheme. The idea is to use equal spacing on the *intrinsic time* λ rather than the external time t (see, however, example 5.4.10 below). Accordingly, fix a step size $\delta > 0$ and set

$$\lambda_k \stackrel{\text{def}}{=} k\delta \quad \text{and} \quad T_k = T_k^{(\delta)} \stackrel{\text{def}}{=} T^{\lambda_k}\ ,\ k = 0, 1, \ldots\ . \tag{5.4.21}$$

This produces a stochastic partition $\mathcal{T} = \mathcal{T}^{(\delta)}$ whose mesh tends to zero as $\delta \to 0$. For our purpose it is convenient to estimate the right-hand side of (5.2.35), which reads

$$\left\| X' - X \right\|^{\star}_{p,M} \leq \frac{\gamma}{L(1-\gamma)} \cdot \left\| \boldsymbol{F}[X'] - \boldsymbol{F}'[X'] \right\|_{p,M}\ .$$

Namely, $\qquad \Delta_t \stackrel{\text{def}}{=} \boldsymbol{F}[X']_t - \boldsymbol{F}'[X']_t = \boldsymbol{f}(X'_t) - \boldsymbol{f}(X'^{\mathcal{T}}_t)$

equals[1] $\qquad \boldsymbol{f}\big(X'_{T_k} + f_\eta(X'_{T_k})\cdot(Z^\eta_t - Z^\eta_{T_k})\big) - \boldsymbol{f}(X'_{T_k})$

for $T_k \leq t < T_{k+1}$, and there satisfies the estimate[35]

$$\begin{aligned} \left|\Delta^\nu_t\right| &\leq L \cdot \left| f^\nu_\eta(X'_{T_k})\cdot(Z^\eta_t - Z^\eta_{T_k}) \right|\ , \qquad \nu = 1, \ldots, n\ , \\ &\leq L \cdot \left| \left\{ \boldsymbol{f}^\nu(X'_{T_k}) (\!(T_k, T_{k+1}]\!] \right\} * \boldsymbol{Z} \right|^{\star}_t\ . \end{aligned}$$

Thus $\qquad \left|\Delta^\nu_t\right| \leq L \cdot \sum_{0\leq k} [\![T_k, T_{k+1})\!)_t \cdot \left| \boldsymbol{f}^\nu(X'_{T_k}) (\!(T_k, T_{k+1}]\!] * \boldsymbol{Z} \right|^{\star}_t$

for *all* $t \geq 0$ and, since THE time transformation is strictly increasing,

$$\begin{aligned} \left\| \Delta^\nu_{T^\mu -} \right\|_{L^p} \leq L \cdot \textstyle\sum_k [k\delta < \mu \leq (k+1)\delta] \times \\ \times \left\| \left| \boldsymbol{f}^\nu(X'_{T_k}) (\!(T_k, T_{k+1}]\!] * \boldsymbol{Z} \right|^{\star}_{T^\mu -} \right\|_{L^p} \end{aligned}$$

(which is a sum with only one non-vanishing term)

$$\text{by (4.5.1) and 2.4.7:} \qquad \leq LC_p^\diamond \cdot \textstyle\sum_k [k\delta < \mu \leq (k+1)\delta] \times$$

$$\times \max_{\rho=1^\diamond, p^\diamond} \left\| \left(\int_{k\delta}^{\mu} \left| \boldsymbol{f}^\nu(X'_{T_k}) \right|_\infty^\rho d\lambda \right)^{1/\rho} \right\|_{L^p}$$

$$\text{for } \delta < 1: \qquad \leq LC_p^\diamond \cdot \textstyle\sum_k [k\delta < \mu \leq (k+1)\delta] \cdot \delta^{1/p^\diamond} \left\| \left| \boldsymbol{f}^\nu(X'_{T_k}) \right|_\infty \right\|_{L^p}.$$

Therefore, applying $|\ |_p$, Fubini's theorem, and inequality (5.4.18),

$$\left\| \Delta_{T^\mu -} \right\|_{L^p} \leq \delta^{1/p^\diamond} \cdot L^2 C_p^\diamond \cdot \left\| X'^\star_{T_k} \right\|_{L^p} \leq \delta^{1/p^\diamond} \cdot L^2 C_p^\diamond \cdot \left\| X'^\star_{T^\mu} \right\|_{L^p} .$$

Multiplying by $e^{-M\mu}$, taking the supremum over μ, and using (5.2.23) results in inequality (5.4.22) below:

Theorem 5.4.9 *Suppose that $\boldsymbol{Z}$ is a quasi-left-continuous local $L^q(\mathbb{P})$-integrator for some $q \geq 2$, let $p \in [2, q]$, $0 < \gamma < 1$, and suppose that the markovian stochastic differential equation (5.4.17) of Lipschitz constant L has its unique global solution in $\mathfrak{S}^{\star n}_{p,M}$, $M = M^{(5.2.26)}_{L:\gamma}$. Then the non-adaptive Euler–Peano approximate X' defined in equation (5.4.19) for $\delta > 0$ satisfies*

$$\left\| X' - X \right\|^\star_{p,M} \leq \delta^{1/p^\diamond} \cdot \frac{C_p^\diamond L\gamma \cdot \left\| {}^0C \right\|^\star_{p,M}}{(1-\gamma)^2} . \tag{5.4.22}$$

Consequently, if δ runs through a sequence δ_n such that $\sum_n \delta_n^{1/p^\diamond}$ converges, then the corresponding non-adaptive Euler–Peano approximates converge uniformly on bounded time-intervals to the exact solution, nearly. ∎

Example 5.4.10 Suppose $\boldsymbol{Z}$ is a Lévy process whose Lévy measure has q^{th} moments away from zero and therefore is an L^q-integrator (see proposition 4.6.16 on page 267). Then its previsible controller is a multiple of time (ibidem), and $T^\lambda = c\lambda$ for some constant c. In that case the classical subdivision into equal time-intervals coincides with the intrinsic one above, and we get the pathwise convergence of the classical Euler–Peano approximates (5.4.19) under the condition $\sum_n \delta_n^{1/p^\diamond} < \infty$. If in particular $\boldsymbol{Z}$ has no jumps and so is a Wiener process, or if $p = 2$ was chosen, then $p^\diamond = 2$, which implies that square-root summability of the sequence of step sizes suffices for pathwise convergence of the non-adaptive Euler–Peano approximates.

Remark 5.4.11 So why not forget the adaptive algorithm (5.4.3)–(5.4.5) and use the non-adaptive scheme (5.4.19) exclusively?

Well, the former algorithm has order[37] 1 (see (5.4.10)), while the latter has only order $1/2$ – or worse if there are jumps, see (5.4.22). (It should be

[37] Roughly speaking, an approximation scheme is of ***order*** r if its ***global error*** is bounded by a multiple of the r^{th} power of the step size. For precise definitions see pages 281 and 324.

pointed out in all fairness that the expected number of computations needed to reach a given final time grows as $1/\delta^2$ in the first algorithm and as $1/\delta$ in the second, when a Wiener process is driving. In other words, the adaptive Euler algorithm essentially has order $1/2$ as well.)

Next, the algorithm (5.4.19) above can (so far) only be shown to make sense and to converge when the driver $\boldsymbol{Z}$ is at least an L^2-integrator. A reduction of the general case to this one by factorization does not seem to offer any practical prospects. Namely, change to another probability in $\mathfrak{P}[\boldsymbol{Z}]$ alters THE time transformation and with it the algorithm: there is no universality property as in summary 5.4.7.

Third, there is the generalization of the adaptive algorithm to general endogenous coupling coefficients (theorem 5.4.5), but not to my knowledge of the non-adaptive one.

The Stratonovich Equation

In this subsection we assume that the drivers Z^η are continuous and the coupling coefficients markovian.[8] On page 271 the original ill-put stochastic differential equation (5.1.1) was replaced by the Itô equation (5.1.2), so as to have its integrands previsible and therefore integrable in the Itô sense. Another approach is to read (5.1.1) as a ***Stratonovich equation***:[38]

$$X = C + f_\eta(X)\circ Z^\eta \stackrel{\text{def}}{=} C + f_\eta(X)*Z^\eta + \frac{1}{2}\big[f_\eta(X), Z^\eta\big] . \tag{5.4.23}$$

Now, in the presence of sufficient smoothness of $\boldsymbol{f}$, there is by Itô's formula a continuous finite variation process V such that[38,16]

$$f_\eta(X) = f_{\eta;\nu}(X)*X^\nu + V .$$

Hence $\big[f_\eta(X), Z^\eta\big] = f_{\eta;\nu}(X)*\big[X^\nu, Z^\eta\big] = f_{\eta;\nu}(X)f_\theta^\nu(X)*\big[Z^\theta, Z^\eta\big]$,

which exhibits the Stratonovich equation (5.4.23) as equivalent with the Itô equation

$$X = C + f_\eta(X)*Z^\eta + \frac{\big(f_{\eta;\nu}f_\theta^\nu\big)(X)}{2}*\big[Z^\theta, Z^\eta\big] \ : \tag{5.4.24}$$

X solves (5.4.23) if and only if it solves (5.4.24). Since the Stratonovich integral has no decent limit properties, the existence and uniqueness of solutions to equation (5.4.23) cannot be established by a contractivity argument. Instead we must read it as the Itô equation (5.4.24); Lipschitz conditions on *both* the f_η and the $f_{\eta;\nu}f_\theta^\nu$ will then produce a unique global solution.

[38] Recall that X, C, f_η take values in $\mathbb{R}^n$. For example, $\boldsymbol{f} = \{\boldsymbol{f}^\nu\} = \{f_\eta^\nu\}$. The indices η, θ, ι usually run from 1 to d and the indices $\mu, \nu, \rho \ldots$ from 1 to n. Einstein's convention, adopted, implies summation over the same indices in opposite positions.

Exercise 5.4.12 (Coordinate Invariance of the Stratonovich Equation) Let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ be invertible and twice continuously differentiable. Set $f_\eta^\Phi(y) \stackrel{\text{def}}{=} \Phi\big(f_\eta(\Phi^{-1}(y))\big)$. Then $Y \stackrel{\text{def}}{=} \Phi(X)$ is the unique solution of

$$Y = \Phi(C) + f_\eta^\Phi(Y)\circ Z^\eta \,,$$

if and only if X is the solution of equation (5.4.23). In other words, the Stratonovich equation behaves like an ordinary differential equation under coordinate transformations – the Itô equation generally does not. This feature, together with theorem 3.9.24 and application 5.4.25, makes the Stratonovich integral very attractive in modeling.

Higher Order Approximation: Obstructions

Approximation schemes of global order $1/2$ as offered in theorem 5.4.9 seem unsatisfactory. From ordinary differential equations we are after all accustomed to Taylor or Runge–Kutta schemes of arbitrarily high order.[37] Let us discuss what might be expected in the stochastic case, at the example of the Stratonovich equation (5.4.23) and its equivalent (5.4.24), reproduced here as

$$X = C + \boldsymbol{f}(X)\circ\boldsymbol{Z} \tag{5.4.25}$$

or[38]

$$X = C + f_\eta(X)*Z^\eta + \frac{\big(f_{\eta;\nu}f_\theta^\nu\big)(X)}{2}*\big[Z^\theta, Z^\eta\big] \tag{5.4.26}$$

or, equivalently,

$$X = \mathfrak{U}[X] \stackrel{\text{def}}{=} C + \overline{F}_\iota(X)*\overline{Z}^\iota \,, \tag{5.4.27}$$

where $\overline{F}_\iota \stackrel{\text{def}}{=} f_\eta$ and $\overline{Z}^\iota \stackrel{\text{def}}{=} Z^\eta$ when $\iota = \eta \in \{1,\ldots,d\}$

and $\overline{F}_\iota \stackrel{\text{def}}{=} f_{\eta;\nu}f_\theta^\nu$ and $\overline{Z}^\iota \stackrel{\text{def}}{=} [Z^\eta, Z^\theta]$ when $\iota = \eta\theta \in \{11,\ldots,dd\}$.

In order to simplify and to fix ideas we work with the following

Assumption 5.4.13 *The initial condition $C \in \mathcal{F}_0$ is constant in time. $\boldsymbol{Z}$ is continuous with $\boldsymbol{Z}_0 = 0$ – then $\overline{\boldsymbol{Z}}_0 = 0$. The markovian[8] coupling coefficient $\overline{\boldsymbol{F}}$ is differentiable and Lipschitz.*

We are then sure that there is a unique solution X of (5.4.27), which also solves (5.4.25) and lies in $\mathfrak{S}_{p,M}^{\star n}$ for any $p \ge 2$ and $M > M_{p,L}^{\diamond(5.2.20)}$ (see proposition 5.2.14).

We want to compare the effect of various step sizes $\delta > 0$ on the accuracy of a given non-adaptive approximation scheme. For every $\delta > 0$ picked, T_k shall denote the ***intrinsically $\boldsymbol{\delta}$-spaced*** stopping times of equation (5.4.21): $T_k \stackrel{\text{def}}{=} T^{k\delta}$.

Surprisingly much – of, alas, a disappointing nature – can be derived from a rather general discussion of single-step approximation methods. We start with the following "metaobservation:" A straightforward[39] generalization of a classical single-step scheme as described on page 280 will result in a method of the following description:

---

[39] and, as it turns out, a bit naive – see notes 5.4.33.

Condition 5.4.14 *The method provides a function* $\Xi' : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$,

$$(x, \boldsymbol{z}) \mapsto \Xi'[x, \boldsymbol{z}] = \Xi'[x, \boldsymbol{z}; \boldsymbol{f}] \,,$$

whose role is this: when after k *steps the method has constructed an approximate solution* X'_t *for times* t *up to the* k^{th} *stopping time* T_k, *then* Ξ' *is employed to extend* X' *up to the next time* T_{k+1} *via*

$$X'_t \stackrel{\text{def}}{=} \Xi'[X'_{T_k}, \boldsymbol{Z}_t - \boldsymbol{Z}^{T_k}] \quad \textit{for} \quad T_k \le t \le T_{k+1} \,. \tag{5.4.28}$$

$\boldsymbol{Z} - \boldsymbol{Z}^{T_k}$ *is the upcoming stretch of the driver. The function* Ξ' *is specific for the method at hand, and is constructed from the coupling coefficient* $\boldsymbol{f}$ *and possibly (in Taylor methods) from a number of its derivatives.*

If the approximation scheme meets this description, then we talk about ***the method*** Ξ'.

In an adaptive[33] scheme, Ξ' might also enter the definition of the next stopping time T_{k+1} – see for instance (5.4.4). The function Ξ' should be reasonably simple; the more complex Ξ' is to evaluate the poorer a choice it is, evidently, for an approximation scheme, unless greatly enhanced accuracy pays for the complexity. In the usual single-step methods $\Xi'[x, \boldsymbol{z}; \boldsymbol{f}]$ is an algebraic expression in various derivatives of $\boldsymbol{f}$ evaluated at algebraic expressions made from x and $\boldsymbol{z}$.

Examples 5.4.15 In the Euler–Peano method of theorem 5.4.9

$$\Xi'[x, \boldsymbol{z}; \boldsymbol{f}] = x + f_\eta(x) z^\eta \,.$$

The classical improved Euler or Heun method generalizes to[1]

$$\Xi'[x, \boldsymbol{z}; \boldsymbol{f}] \stackrel{\text{def}}{=} x + \frac{f_\eta(x) + f_\eta(x + f_\theta(x) z^\theta)}{2} z^\eta \,.$$

The straightforward[39] generalization of the Taylor method of order 2 is given by

$$\Xi'[x, \boldsymbol{z}; \boldsymbol{f}] \stackrel{\text{def}}{=} x + f_\eta(x) z^\eta + \big(f_{\eta;\nu} f_\theta^\nu\big)(x) z^\eta z^\theta / 2 \,.$$

The classical Runge–Kutta method of global order 4 has the obvious generalization

$$k_1 \stackrel{\text{def}}{=} f_\eta(x) z^\eta \,,\ k_2 \stackrel{\text{def}}{=} f_\eta(x + k_1/2) z^\eta \,,\ k_3 \stackrel{\text{def}}{=} f_\eta(x + k_2/2) z^\eta \,,\ k_4 \stackrel{\text{def}}{=} f_\eta(x + k_3/2) z^\eta$$

and
$$\Xi'[x, \boldsymbol{z}; \boldsymbol{f}] \stackrel{\text{def}}{=} x + \frac{k_1 + 2k_2 + 2k_3 + k_4}{6} \,.$$

The methods Ξ' in this example have a structure in common that is most easily discussed in terms of the following notion. Let us say that the map $\Phi : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$ is ***polynomially bounded in*** $\boldsymbol{z}$ if there is a polynomial P so that

$$|\Phi[x, \boldsymbol{z}]| \le P(|\boldsymbol{z}|) \,, \qquad (x, \boldsymbol{z}) \in \mathbb{R}^n \times \mathbb{R}^d \,.$$

The functions polynomially bounded in $\boldsymbol{z}$ evidently form an algebra $\mathcal{BP}$ that is closed under composition: $\Psi\big[\Phi[\,\cdot\,,\cdot\,],\cdot\,\big] \in \mathcal{BP}$ for $\Phi, \Psi \in \mathcal{BP}$. The functions $\Phi \in \mathcal{BP} \cap C^k$ whose first k partials belong to $\mathcal{BP}$ as well form the class $\mathcal{PU}^k$. This is easily seen to form again an algebra closed under composition. For simplicity's sake assume now that $\boldsymbol{f}$ is of class[10] C_b^∞. Then in the examples above, and in fact in all straightforward extensions of the classical single step methods, $\Xi'[\,\cdot\,,\cdot\,;\boldsymbol{f}]$ has all of its partial derivatives in $\mathcal{BP}^\infty$. For the following discussion only this much is needed:

Condition 5.4.16 *Ξ' has partial derivatives of orders 1 and 2 in $\mathcal{BP}$.*

Now, from definition (5.4.28) on page 322 and theorem 3.9.24 on page 170 we get $\Xi'[x,0]=x$ and[40]

$$\Xi'[X_{T_k},\boldsymbol{Z}-\boldsymbol{Z}^{T_k}] = X_{T_k} + \Xi'_{;\eta}[X_{T_k},\boldsymbol{Z}-\boldsymbol{Z}^{T_k}]{\circ}Z^\eta \quad \text{on} \quad [\![T_k,T_{k+1}]\!] \,,$$

so that X' can be viewed as the solution of the Stratonovich equation

$$X' = C + F'_\eta[X']{\circ}Z^\eta \,, \tag{5.4.29}$$

with Itô equivalent (compare with equation (5.4.27) on page 321)

$$X' = \mathfrak{U}'[X'] \stackrel{\text{def}}{=} C + \overline{F}'_\iota[X']{*}\overline{Z}^\iota \,, \tag{5.4.30}$$

where[35]
$$\overline{F}'_\iota = F'_\eta \stackrel{\text{def}}{=} \textstyle\sum_k [\![T_k,T_{k+1})\!) \cdot \Xi'_{;\eta}[X'_{T_k},\boldsymbol{Z}-\boldsymbol{Z}^{T_k}] \qquad \text{when } \iota=\eta$$

and
$$\overline{F}'_\iota \stackrel{\text{def}}{=} \textstyle\sum_k [\![T_k,T_{k+1})\!) \cdot \big\{\Xi'_{;\eta\nu}\Xi'^{\nu}_{;\theta}\big\}[X'_{T_k},\boldsymbol{Z}-\boldsymbol{Z}^{T_k}] \qquad \text{for } \iota=\eta\theta \,.$$

Note that $\overline{\boldsymbol{F}}'$ is generally not markovian, in view of the explicit presence of $\boldsymbol{Z}-\boldsymbol{Z}^{T_k}$ in $\Xi'_{;\eta}[x,\boldsymbol{Z}-\boldsymbol{Z}^{T_k}]$.

Exercise 5.4.17 (i) Condition 5.4.16 ensures that $\overline{\boldsymbol{F}}'$ satisfies the Lipschitz condition (5.2.11). Therefore both maps $\mathfrak{U}$ of (5.4.27) and $\mathfrak{U}'$ of (5.4.30) will be strictly contractive in $\mathfrak{S}^{\star n}_{p,M}$ for all $p \ge 2$ and suitably large $M=M(p)$. (ii) Furthermore, there exist constants D', L', M' such that for $0 \le \kappa < \lambda$

$$\Big\| |\Xi'[C,\boldsymbol{Z}_{\boldsymbol{.}}-\boldsymbol{Z}^{T^\kappa};\boldsymbol{f}]|^\star_{T^\lambda} \Big\|_{L^p} \le D' \cdot \|C\|_{L^p} \cdot e^{M'(\lambda-\kappa)} \tag{5.4.31}$$

and
$$\Big\| |\Xi'[C',\boldsymbol{Z}_{\boldsymbol{.}}-\boldsymbol{Z}^{T^\kappa}] - \Xi'[C,\boldsymbol{Z}_{\boldsymbol{.}}-\boldsymbol{Z}^{T^\kappa}]|^\star_{T^\lambda} \Big\|_{L^p} \le \big\|C'-C\big\|_{L^p} \cdot e^{L'(\lambda-\kappa)} \,. \tag{5.4.32}$$

Recall that we are after a method Ξ' of order strictly larger than $1/2$. That is to say, we want it to produce an estimate of the form $X'-X=o(\sqrt{\delta})$ for the difference of the exact solution $X=\Xi[C,\boldsymbol{Z};\boldsymbol{f}]$ of (5.4.25) from its Ξ'-approximate X' made with step size δ via (5.4.28). The question arises how to measure this difference. We opt[41] for a generalization of the classical notions of order from page 281, replacing time t with intrinsic time λ:

[40] We write $\Xi'_{;\eta} \stackrel{\text{def}}{=} \partial\Xi'/\partial z^\eta$ and $\Xi'_{;\eta\nu} \stackrel{\text{def}}{=} \partial\Xi'_{;\eta}/\partial x^\nu$, etc.[38]

[41] There are less stringent notions; see notes 5.4.33.

Definition 5.4.18 *We say that* Ξ' *has* ***local order*** r *on the coupling coefficient* $\boldsymbol{f}$ *if there exists a constant* $\underline{M}$ *such that*[4] *for all* $\lambda > \kappa \geq 0$ *and all* $C \in L^p(\mathcal{F}_{T^\lambda})$

$$\left\| \left| \Xi'[C, \boldsymbol{Z}_{\boldsymbol{.}} - \boldsymbol{Z}^{T^\kappa}; \boldsymbol{f}] - \Xi[C, \boldsymbol{Z}_{\boldsymbol{.}} - \boldsymbol{Z}^{T^\kappa}; \boldsymbol{f}] \right|^\star_{T^\lambda} \right\|_{L^p} \tag{5.4.33}$$
$$\leq (\|C\|_{L^p}+1) \times \big(\underline{M}(\lambda-\kappa)\big)^r e^{\underline{M}(\lambda-\kappa)} .$$

The least such $\underline{M}$ *is denoted by* $\underline{M}[\boldsymbol{f}]$. *We say* Ξ' *has* ***global order*** r *on* $\boldsymbol{f}$ *if the difference* $X' - X$ *satisfies an estimate*

$$\big\| X'_{\boldsymbol{.}} - X_{\boldsymbol{.}} \big\|^\star_{p,\overline{M}} = (\big\| C \big\|^\star_{p,\overline{M}} + 1) \cdot O(\delta^r)$$

for some $\overline{M} = \overline{M}[\boldsymbol{f}; \Xi']$. *This amounts to the existence of a* $\overline{B} = \overline{B}[\boldsymbol{f}; \Xi']$

such that $$\left\| \left| X' - \Xi[C, \boldsymbol{Z}; \boldsymbol{f}] \right|^\star_{T^\lambda} \right\|_{L^p} \leq \overline{B} \cdot (\|C\|_{L^p}+1) \times \delta^r e^{\overline{M}\lambda} \tag{5.4.34}$$

for sufficiently small $\delta > 0$ *and all* $\lambda \geq 0$ *and* $C \in L^p(\mathcal{F}_0)$.

Criterion 5.4.19 *(Compare with criterion 5.1.11.) Assume condition 5.4.16.*

(i) If $\left\| \left| \Xi'[C, \boldsymbol{Z}_{\boldsymbol{.}} - \boldsymbol{Z}^{T^\kappa}; \boldsymbol{f}] - \Xi[C, \boldsymbol{Z}_{\boldsymbol{.}} - \boldsymbol{Z}^{T^\kappa}; \boldsymbol{f}] \right|^\star_{T^\lambda} \right\|_{L^p} = (\|C\|_{L^p}+1) \cdot O((\lambda-\kappa)^r)$, *then* Ξ' *has local order* r *on* $\boldsymbol{f}$.

(ii) If Ξ' *has local order* r, *then it has global order* $r-1$.

Recall again that we are after a method Ξ' of order strictly larger than $1/2$. In other words, we want it to produce

$$\big\| X' - X \big\|^\star_{p,\overline{M}} = o(\sqrt{\delta}) \tag{5.4.35}$$

for some $p \geq 2$ and some $\overline{M}$. Let us write ${}^0\Xi'(t) \stackrel{\text{def}}{=} \Xi'[C, \boldsymbol{Z}_t] - C$ and ${}^0\Xi'_{;\eta}(t) \stackrel{\text{def}}{=} \Xi'_{;\eta}[C, \boldsymbol{Z}_t]$ for short.[40] According to inequality (5.2.35) on page 294,

(5.4.35) will follow from $$\big\| \overline{\boldsymbol{F}}[X'] - \overline{\boldsymbol{F}}'[X'] \big\|_{p,\overline{M}} = o(\sqrt{\delta}) ,$$

which requires $$\big\| f_\eta\big(C + {}^0\Xi'(t)\big) - {}^0\Xi'_{;\eta}(t) \big\|_{L^p} = o(\sqrt{t}) . \tag{5.4.36}$$

It is hard to see how (5.4.35) could hold without (5.4.36); at the same time, it is also hard to establish that it implies (5.4.36). We will content ourselves with this much:

Exercise 5.4.20 If Ξ' is to have order > 1 in all circumstances, in particular whenever the driver $\boldsymbol{Z}$ is a standard Wiener process, then equation (5.4.36) must hold.

Letting $\delta \to 0$ in (5.4.36) we see that the method Ξ' must satisfy $\Xi'_{;\eta}[C, 0] = f_\eta(C)$. This can be had in all generality only if[40]

$$\Xi'_{;\eta}[x, 0] = f_\eta(x) \quad \forall\, x \in \mathbb{R}^n . \tag{5.4.37}$$

Then[5]
$$f_\eta(C + {}^0\Xi'(t)) = f_\eta(C) + f_{\eta;\nu}(C)\, {}^0\Xi'^\nu(t) + O(|{}^0\Xi'(t)|^2)$$
$$= f_\eta(C) + f_{\eta;\nu}(C)\, \Xi'^\nu_{;\theta}[C,0] Z^\theta_t + O(|\boldsymbol{Z}_t|^2)$$
by (5.4.37):
$$= f_\eta(C) + f_{\eta;\nu}(C) f^\nu_\theta(C)\, Z^\theta_t + O(|\boldsymbol{Z}_t|^2) \; . \tag{5.4.38}$$

Also,
$$\Xi'_{;\eta}[C, \boldsymbol{Z}_t] = f_\eta(C) + \Xi'_{;\eta\theta}[C,0] Z^\theta_t + O(|\boldsymbol{Z}_t|^2) \; . \tag{5.4.39}$$

Equations (5.4.36), (5.4.38), and (5.4.39) imply that for $t \le T^\delta$
$$\Big\| \Big\{ \big(f_{\eta;\nu} f^\nu_\theta\big)(C) - \Xi'_{;\eta\theta}[C,0] \Big\} Z^\theta_t \Big\|_{L^p} = o(\sqrt{\delta}) + \big\| O(|\boldsymbol{Z}_t|^2) \big\|_{L^p} \; . \tag{5.4.40}$$

This condition on Ξ' can be had, of course, if Ξ' is chosen so that
$$M_{\eta\theta}(x) \stackrel{\text{def}}{=} \big(f_{\eta;\nu} f^\nu_\theta\big)(x) - \Xi'_{;\eta\theta}[x,0] = 0 \quad \forall\, x \in \mathbb{R}^n \; , \tag{5.4.41}$$

and in general only with this choice. Namely, suppose $\boldsymbol{Z}$ is a standard d-dimensional Wiener process. Then, for $k = 0$, the size in L^p of the martingale $M_t \stackrel{\text{def}}{=} M^\mu_{\eta\theta}(x) Z^\theta_t$ at $t = \delta$ is, by theorem 2.5.19 and inequality (4.2.4), bounded below by a multiple of
$$\|S_\delta[M_.]\|_{L^p} = \Big\| \Big(\textstyle\sum_\theta \big| M^\mu_{\eta\theta}(x) \big|^2 \Big)^{1/2} \Big\|_{L^p} \cdot \sqrt{\delta} \; ,$$
while
$$\big\| O(|\boldsymbol{Z}_\delta|^2) \big\|_{L^p} \le const \times \delta = o(\sqrt{\delta}) \; .$$

In the presence of equation (5.4.40), therefore,
$$\Big\| \Big(\textstyle\sum_\theta \big| M^\mu_{\eta\theta}(x) \big|^2 \Big)^{1/2} \Big\|_{L^p} \le \frac{o(\sqrt{\delta})}{\sqrt{\delta}} \xrightarrow[\delta \to 0]{} 0 \; .$$

This implies $M_. = 0$ and with it (5.4.41), i.e., $\Xi'_{;\eta\theta}[x,0] = \big(f_{\eta;\nu} f^\nu_\theta\big)(x)$ for all $x \in \mathbb{R}^n$. Notice now that $\Xi'_{;\eta\theta}[x,0]$ is symmetric in η, θ. This equality therefore implies that the Lie brackets $[f_\eta, f_\theta] \stackrel{\text{def}}{=} f_{\eta;\nu} f^\nu_\theta - f_{\theta;\nu} f^\nu_\eta$ must vanish:

Condition 5.4.21 *The vector fields $f_1, \ldots, f_d$ commute.*

The following summary of these arguments does not quite deserve to be called a theorem, since the definition of a method and the choice of the norms $\|\;\|_{p,M}$, etc., are not canonical and (5.4.36) was not established rigorously.

Scholium 5.4.22 *We cannot expect a method Ξ' satisfying conditions 5.4.14 and 5.4.16 to provide approximation in the sense of definition 5.4.18 to an order strictly better than $1/2$ for all drivers and all initial conditions, unless the coefficient vector fields commute.*

Higher Order Approximation: Results

We seek approximation schemes of an order better than $1/2$. We continue to investigate the Stratonovich equation (5.4.25) under assumption 5.4.13, adding condition 5.4.21. This condition, forced by scholium 5.4.22, is a severe restriction on the system (5.4.25). The least one might expect in a just world is that in its presence there are good approximation schemes. Are there? In a certain sense, the answer is affirmative and optimal. Namely, from the change-of-variable formula (3.9.11) on page 171 for the Stratonovich integral, this much is immediate:

Theorem 5.4.23 *Assuming condition 5.4.21, let $\Xi^{\boldsymbol{f}}$ be the action of $\mathbb{R}^d$ on $\mathbb{R}^n$ generated by $\boldsymbol{f}$ (see proposition 5.1.10 on page 279). Then the solution of equation (5.4.25) is given by*

$$X_t = \Xi^{\boldsymbol{f}}[C, \boldsymbol{Z}_t] .$$

Examples 5.4.24 (i) Let W be a standard Wiener process. The Stratonovich equation $\mathscr{E} = 1 + \mathscr{E}\circ W$ has the solution e^W, on the grounds that $e^{\boldsymbol{\cdot}}$ solves the corresponding ordinary differential equation $e^t = 1 + \int_0^t e^s\, ds$.

(ii) The vector fields $f_1(x) = x$ and $f_2(x) = -x/2$ on $\mathbb{R}$ commute. Their flows are $\xi[x,t;f_1] = xe^t$ and $\xi[x,t;f_2] = xe^{-t/2}$, respectively, and so the action $\boldsymbol{f} = (f_1, f_2)$ generates is $\Xi^{\boldsymbol{f}}(x,(z_1,z_2)) = x \times e^{z_1} \times e^{-z_2/2}$. Therefore the solution of the Itô equation $\mathscr{E}_t = 1 + \int_0^t \mathscr{E}_s\, dW_s$, which is the same as the Stratonovich equation $\mathscr{E}_t = 1 + \int_0^t \mathscr{E}_s\, \delta W_s - 1/2 \int_0^t \mathscr{E}_s\, ds = 1 + \int_0^t f_1(\mathscr{E}_s)\, \delta W_s + \int_0^t f_2(\mathscr{E}_s)\, ds$, is $\mathscr{E}_t = e^{W_t - t/2}$, which the reader recognizes from proposition 3.9.2 as the Doléans–Dade exponential of W.

(iii) The previous example is about linear stochastic differential equations. It has the following generalization. Suppose $A_1, \ldots, A_d$ are *commuting* $n \times n$-matrices. The vector fields $f_\eta(x) \stackrel{\text{def}}{=} A_\eta x$ then commute. The linear Stratonovich equation $X = C + A_\eta X \circ Z^\eta$ then has the explicit solution $X_t = C \cdot e^{A_\eta Z_t^\eta}$. The corresponding Itô equation $X = C + A_\eta X * Z^\eta$, equivalent with $X = C + A_\eta X \circ Z^\eta - \frac{1}{2} A_\eta A_\theta X \circ [Z^\eta, Z^\theta]$, is solved explicitely by $X_t = C \cdot e^{A_\eta Z_t^\eta - \frac{1}{2} A_\eta A_\theta [Z^\eta, Z^\theta]_t}$.

Application 5.4.25 (Approximating the Stratonovich Equation by an ODE) Let us continue the assumptions of theorem 5.4.23. For $n \in \mathbb{N}$ let $\boldsymbol{Z}^{(n)}$ be that continuous and piecewise linear process which at the times k/n equals $\boldsymbol{Z}$, $k = 1, 2, \ldots$. Then $\boldsymbol{Z}^{(n)}$ has finite variation but is generally not adapted; the solution of the ordinary differential equation $X_t^{(n)} = C + \int_0^t \boldsymbol{f}\big(X_s^{(n)}\big)\, d\boldsymbol{Z}_s^{(n)}$ (which depends of course on the parameter $\omega \in \Omega$) converges uniformly on bounded time-intervals to the solution X of the Stratonovich equation (5.4.25), for every $\omega \in \Omega$. This is simply because $\boldsymbol{Z}_t^{(n)}(\omega) \to \boldsymbol{Z}_t(\omega)$ uniformly on bounded intervals and $X_t^{(n)}(\omega) = \Xi^{\boldsymbol{f}}\big[C, \boldsymbol{Z}_t^{(n)}(\omega)\big]$. This feature, together with theorem 3.9.24 and exercise 5.4.12, makes the Stratonovich integral very attractive in modeling. ∎

One way of reading theorem 5.4.23 is that $(x, \boldsymbol{z}) \mapsto \Xi^{\boldsymbol{f}}[x, \boldsymbol{z}]$ is a method of infinite order: there is no error. Another, that in order to solve the stochastic differential equation (5.4.25), one merely needs to solve d ordinary differential equations, producing $\Xi^{\boldsymbol{f}}$, and then evaluate $\Xi^{\boldsymbol{f}}$ at $\boldsymbol{Z}$. All of this looks very satisfactory, until one realizes that $\Xi^{\boldsymbol{f}}$ is not at all a simple function to evaluate and that it does not lend itself to run time approximation of X.

5.4.26 A Method of Order r An obvious remedy leaps to the mind: approximate the action $\Xi^{\boldsymbol{f}}$ by some less complex function Ξ'; then $\Xi'[x, \boldsymbol{Z}_t]$ should be an approximation of X_t. This simple idea can in fact be made to work. For starters, observe that one needs to solve *only one* ordinary differential equation in order to compute $X_t(\omega) = \Xi^{\boldsymbol{f}}[C(\omega), \boldsymbol{Z}_t(\omega)]$ for any given $\omega \in \Omega$. Indeed, by proposition 5.1.10 (iii), $X_t(\omega)$ is the value x_{τ_t} at $\tau_t \overset{\text{def}}{=} |\boldsymbol{Z}_t(\omega)|$ of the solution $x_{\bullet}$ to the ODE

$$x_{\bullet} = C(\omega) + \int_0^{\bullet} f(x_\sigma)\, d\sigma \quad \text{, where} \quad f(x) \overset{\text{def}}{=} \textstyle\sum_\eta f_\eta(x)\, Z_t^\eta(\omega)/\tau_t\ . \tag{5.4.42}$$

Note that knowledge of the whole path of $\boldsymbol{Z}$ is not needed, only of its value $\boldsymbol{Z}_t(\omega)$. We may now use any classical method to approximate $x_{\tau_t} = X_t(\omega)$. Here is a suggestion: given an r, choose a classical method ξ' of global order r, for instance a suitable Runge–Kutta or Taylor method, and use it with step size δ to produce an approximate solution $x'_{\bullet} = x'_{\bullet}[c; \delta, f]$ to (5.4.42). According to page 324, to say that the method ξ' chosen has global order r means that there are constants $\overline{b} = \overline{b}[c; f, \xi']$ and $\overline{m} = \overline{m}[f; \xi']$ so that for sufficiently small $\delta > 0$

$$|x_\sigma - x'_\sigma| \le \overline{b}\cdot\delta^r \times e^{\overline{m}\sigma}\ , \qquad \sigma \ge 0\ .$$

Now set
$$b \overset{\text{def}}{=} \sup\{\overline{b}[f_\eta z^\eta; \xi'] : |\boldsymbol{z}| \le 1\} \tag{5.4.43}$$

and
$$m \overset{\text{def}}{=} \sup\{\overline{m}[f_\eta z^\eta; \xi'] : |\boldsymbol{z}| \le 1\}\ . \tag{5.4.44}$$

Then
$$\big|X_t(\omega) - x'_{\tau_t}\big| \le b\cdot\delta^r \times e^{m\tau_t}\ . \tag{5.4.45}$$

Hidden in (5.4.43), (5.4.44) is another assumption on the method ξ':

Condition 5.4.27 *If ξ' has global order r on $f_1, \ldots, f_d$, then the suprema in equations (5.4.43) and (5.4.44) can be had finite.* (If, as is often the case, $\overline{b}[f; \xi']$ and $\overline{m}[f; \xi']$ can be estimated by polynomials in the uniform bounds of various derivatives of f, then the present condition is easily verified.) ❏

In order to match (5.4.47) with our general definition 5.4.14 of a single-step method, let us define the function $\Xi' : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$ by

$$\Xi'[x, z] \overset{\text{def}}{=} x'_\tau\ , \tag{5.4.46}$$

where $\tau \overset{\text{def}}{=} |\boldsymbol{z}|$ and $x'_{\bullet}$ is the ξ'-approximate to $x_{\bullet} = x + \int_0^{\bullet} f_\eta(x_\sigma) z^\eta/\tau\, d\sigma$. Then the corresponding Ξ'-approximate for (5.4.25) is

$$X'_t(\omega) = \Xi'[c, \boldsymbol{Z}_t(\omega)] \overset{\text{def}}{=} x'_{\tau_t(\omega)}$$

and by (5.4.45) has
$$\big|X(\omega) - X'(\omega)\big|_t^\star \le b\cdot\delta^r \times e^{m|\boldsymbol{Z}|_t^\star(\omega)} \tag{5.4.47}$$

when ξ' is carried out with step size δ. The method Ξ' is still not very simple, requiring as it does $\lceil \tau_t(\omega)/\delta \rceil$ iterations of the classical method ξ' that defines it; but given today's fast computers, one might be able to live with this much complexity. Here is another mitigating observation: if one is interested only in approximating $X_t(\omega)$ at one finite time t, then Ξ' is actually evaluated only once: it is a *one-single-step method.*

Suppose now that $\boldsymbol{Z}$ is in particular of the following ubiquitous form:

Condition 5.4.28 $$Z_t^\eta = \begin{cases} t & \textit{for } \eta = 1 \\ W_t^\eta & \textit{for } \eta = 2, \ldots, d, \end{cases}$$

where $\boldsymbol{W}$ is a standard $d{-}1$-dimensional Wiener process.

Then the previsible controller becomes $\Lambda_t = d{\cdot}t$ (exercise 4.5.19), THE time transformation is given by $T^\lambda = \lambda/d$, and the Stratonovich equation (5.4.25) reads

$$X = C + \boldsymbol{f}(X){\circ}\boldsymbol{Z} \tag{5.4.48}$$

or, equivalently, $$X = C + \overline{f}_\eta(X){*}Z^\eta \ ,$$

where $$\overline{f}_\eta \overset{\text{def}}{=} \begin{cases} f_1 + \dfrac{1}{2} \displaystyle\sum_{\theta>1} f_{\theta;\nu} f_\theta^\nu & \text{for } \eta = 1, \\ f_\eta & \text{for } \eta > 1. \end{cases}$$

$$\sup_{\eta \ge 1} \left| \overline{f}_\eta(x) - \overline{f}_\eta(y) \right| \le L \cdot |x - y| \tag{5.4.49}$$

is the requisite Lipschitz condition from 5.4.13, which guarantees the existence of a unique solution to (5.4.48), which lies in $\mathfrak{S}^{\star n}_{p,M}$ for any $p \ge 2$ and $M > M^{\diamond(5.2.20)}_{p,L}$. Furthermore, $\boldsymbol{Z}$ is of the form discussed in exercise 5.2.18 (ii) on page 292, and inequality (5.4.47) together with inequality (5.2.30) leads to the existence of constants $B' = B'[b, d, p, r]$ and $M' = M'[d, m, p, r]$ such that

$$\left\| \, |X' - X|_t^\star \, \right\|_{L^p} \le \delta^r \cdot B' e^{M't} \ , \qquad t \ge 0 \ .$$

We have established the following result:

Proposition 5.4.29 *Suppose that the driver $\boldsymbol{Z}$ satisfies condition 5.4.28, the coefficients $\overline{f}_1, \ldots, \overline{f}_d$ are Lipschitz, and the coefficients $f_1, \ldots, f_d$ commute. If ξ' is any classical single-step approximation method of global order r for ordinary differential equations in $\mathbb{R}^n$ (page 280) that satisfies condition 5.4.27, then the one-single-step method Ξ' defined from it in (5.4.46) is again of global order r, in this weak sense: at any fixed time t the difference of the exact solution $X_t = \Xi[C, \boldsymbol{Z}_t]$ of (5.4.25) and its Ξ'-approximate X'_t made with step size δ can be estimated as follows: there exist constants B, M, B_1, M_1 that depend only on $d, \boldsymbol{f}, p > 1, \xi'$ such that*

$$\left| X'_t(\omega) - X_t(\omega)) \right| \le B{\cdot}\delta^r \times e^{M|\boldsymbol{Z}_t(\omega)|} \qquad \forall\, \omega \in \Omega$$

and $$\left\| |X' - X|_t^\star \right\|_{L^p} \le B_1{\cdot}\delta^r \times e^{M_1 t} \ . \tag{5.4.50}$$

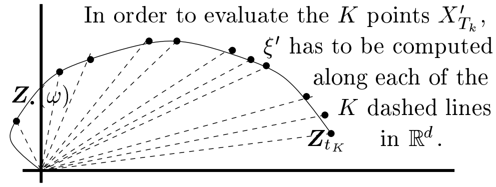

Figure 5.13

Discussion 5.4.30 This result apparently has two related shortcomings: the method Ξ' computes an approximation to the value of $X_t(\omega)$ only at the final time t of interest, not to the whole path $X_{\bullet}(\omega)$, and it waits until that time before commencing the computation – no information from the signal $\boldsymbol{Z}_{\bullet}$ is processed until the final time t has arrived. In order to approximate K points on the solution path $X_{\bullet}$ the method Ξ' has to be run K times, each time using $|\boldsymbol{Z}_{T_k}|/\delta$ iterations of the classical method ξ'. In the figure above[42] one expects to perform as many calculations as there are dashes, in order to compute approximations at the K dots.

Exercise 5.4.31 Suppose one wants to compute approximations $X'_{k\delta}$ at the K points $\delta, 2\delta, \ldots, K\delta = t$ via proposition 5.4.29. Then the expected number of evaluations of ξ' is $N_1 \approx B_1(t)/\delta^2$; in terms of the mean error $E \overset{\text{def}}{=} \| |X' - X|_t^\star \|_{L^2}$

$$N_1 \approx C_1(t)/E^{2/r} ,$$

$B_1(t)$, $C_1(t)$ being functions of at most exponential growth that depend only on ξ'.

Figure 5.13 suggests that one should look for a method that at the $k+1^{\text{th}}$ step uses the previous computations, or at least the previously computed value X'_{T_k}. The simplest thing to do here is evidently to apply the classical method ξ' at the k^{th} point to the ordinary differential equation

$$x_\tau = X'_{T_k} + \int_{T_k}^{\tau} \big(\boldsymbol{f}\cdot(\boldsymbol{Z}_t - \boldsymbol{Z}_t^{T_k})\big)(x_\sigma)\, d\sigma ,$$

whose exact solution at $\tau = 1$ is $x_1 = \Xi^{\boldsymbol{f}}[X'_{T_k}, \boldsymbol{Z}_t - \boldsymbol{Z}_t^{T_k}]$, so as to obtain $X'_t \overset{\text{def}}{=} x'_1$; apply it in one "giant" step of size 1. In figure 5.13 this propels us from one dot to the next. This prescription defines a single-step method Ξ' in the sense of 5.4.14:

$$\Xi'[x, \boldsymbol{z}; \boldsymbol{f}] \overset{\text{def}}{=} \xi'[x, 1; \boldsymbol{f}\cdot\boldsymbol{z}] ;$$

and $\quad X'_t = \Xi'[X'_{T_k}, \boldsymbol{Z}_t - \boldsymbol{Z}_t^{T_k}; \boldsymbol{f}] = \xi'[X'_{T_k}, 1; \boldsymbol{f}\cdot(\boldsymbol{Z}_t - \boldsymbol{Z}_t^{T_k})] , \quad T_k \le t \le T_{k+1} ,$

is the corresponding approximate as in definition (5.4.28).

Exercise 5.4.32 Continue to consider the Stratonovich equation (5.4.48), assuming conditions 5.4.21, 5.4.28, and inequality (5.4.49). Assume that the classical method ξ' is *scale-invariant* (see note 5.1.12) and has local order $r+1$ – by criterion 5.1.11 on page 281 it has global order r. Show that then Ξ' has global order $r/2 - 1/2$ in the sense of (5.4.34), so that, for suitable constants B_2, M_2, the Ξ'-approximate X' satisfies

$$E \overset{\text{def}}{=} \big\| |X' - X|_t^\star \big\|_{L^2} \le B_2(t)\delta^{r/2-1/2} .$$

Consequently the number $N_2 = t/\delta$ of evaluations of ξ' needed as in 5.4.31 is

$$N_2 \approx C_2(t)/E^{2/(r-1)} .$$

[42] It is highly stylized, not showing the wild gyrations the path $\boldsymbol{Z}_{\bullet}$ will usually perform.

In order to decrease the error E by a factor of $10^{r/2}$, we have to increase the expected number of evaluations of the method ξ' by a factor of 10 in the procedure of exercise 5.4.31. The number of evaluations increases by a factor of $10^{r/r-1}$ using exercise 5.4.32 with the estimate given there. We see to our surprise that the procedure of exercise 5.4.31 is better than that of exercise 5.4.32, at least according to the estimates we were able to establish.

Notes 5.4.33 (i) The adaptive Euler method of theorem 5.4.2 is from [7]. It, its generalization 5.4.5, and its non-adaptive version 5.4.9 have global order $1/2$ in the sense of definition 5.4.18. Protter and Talay show in [93] that the latter method has order 1 when the driver is a suitable Lévy process, the coupling coefficients are suitably smooth, and the deviation of the approximate X' from the exact solution X is measured by $\mathbb{E}[g \circ X_t - g \circ X'_t]$ for suitably (rather) smooth g.

(ii) That the coupling coefficients should commute surely is rare. The reaction to scholium 5.4.22 nevertheless should not be despair. Rather, we might distance ourselves from the definition 5.4.14 of a method and possibly entertain less stringent definitions of order than the one adopted in definition 5.4.18. We refer the reader to [86] and [87].

5.5 Weak Solutions

Example 5.5.1 (Tanaka) Let W be a standard Wiener process on its own natural filtration $\mathcal{F}_{.}[W]$, and consider the stochastic differential equation

$$X = \mathrm{sign} X {*} W \ . \tag{5.5.1}$$

The coupling coefficient $$\mathrm{sign} x \stackrel{\mathrm{def}}{=} \begin{cases} 1 & \text{for } x \ge 0 \\ -1 & \text{for } x < 0 \end{cases}$$

is of course more general than the ones contemplated so far; it is, in particular, not Lipschitz and returns a previsible rather than a left-continuous process upon being fed $X \in \mathfrak{C}$. Let us show that (5.5.1) cannot have a strong solution in the sense of page 273. By way of contradiction assume that X solves this equation. Then X is a continuous martingale with square function $[X,X]_t = \Lambda_t \stackrel{\mathrm{def}}{=} t$ and $X_0 = 0$, so it is a standard Wiener process (corollary 3.9.5). Then

$$|X|^2 = X^2 = 2X{*}X + \Lambda = 2X \mathrm{sign} X{*}W + \Lambda \ ,$$

and so $$\frac{1}{|X|+\epsilon}{*}|X|^2 = \frac{2|X|}{|X|+\epsilon}{*}W + \frac{1}{|X|+\epsilon}{*}\Lambda \ , \qquad \epsilon > 0 \ .$$

Then $$W = \lim_{\epsilon \to 0} \frac{|X|}{|X|+\epsilon}{*}W = \lim_{\epsilon \to 0} \frac{1}{|X|+\epsilon}{*}(|X|^2 - \Lambda)/2$$

is adapted to the filtration generated by $|X|$: $\mathcal{F}_t[X] \subseteq \mathcal{F}_t[W] \subseteq \mathcal{F}_t[|X|] \quad \forall t$ – this would make X a Wiener process adapted to the filtration generated by its absolute value $|X|$, what nonsense. Thus (5.5.1) has no strong solution. Yet it has a solution in some sense: start with a Wiener process X on its own natural filtration $\mathcal{F}_.[X]$, and set $W \stackrel{\text{def}}{=} \operatorname{sign} X * X$. Again by corollary 3.9.5, W is a standard Wiener process on $\mathcal{F}_.[X]$ (!), and equation (5.5.1) is satisfied. In fact, there is more than one solution, $-X$ being another one. What is going on? In short: the natural filtration of the driver W of (5.5.1) was too small to sustain a solution of (5.5.1).

Example 5.5.1 gives rise to the notion of a weak solution. To set the stage consider the stochastic differential equation

$$X = C + f_\eta[\boldsymbol{Z}, X]_{.-} * Z^\eta = C + \boldsymbol{f}[\boldsymbol{Z}, X]_{.-} * \boldsymbol{Z} . \tag{5.5.2}$$

Here $\boldsymbol{Z}$ is our usual vector of integrators on a measured filtration $(\mathcal{F}_., \mathbb{P})$. The coupling coefficients f_η are assumed to be endogenous and to act in a non-anticipating fashion – see (5.1.4):

$$f_\eta\big[\boldsymbol{z}_., x_.\big]_t = f_\eta\big[\boldsymbol{z}_.^t, x_.^t\big]_t \qquad \forall\, \boldsymbol{z}_. \in \mathscr{D}^d,\ x_. \in \mathscr{D}^n,\ t \geq 0 .$$

Definition 5.5.2 *A* ***weak solution*** Ξ' *of equation (5.5.2) is a filtered probability space* $(\Omega', \mathcal{F}'_., \mathbb{P}')$ *together with* $\mathcal{F}'_.$*-adapted processes* $C', \boldsymbol{Z}', X'$ *such that the law of* $(C', \boldsymbol{Z}')$ *on* $\mathscr{D}^{n+d}$ *is the same as that of* $(C, \boldsymbol{Z})$*, and such that (5.5.2) is satisfied:*

$$X' = C' + \boldsymbol{f}[\boldsymbol{Z}', X']_{.-} * \boldsymbol{Z}' .$$

The problem (5.5.2) is said to have a ***unique weak solution*** *if for any other weak solution* $\Xi'' = (\Omega'', \mathcal{F}''_., \mathbb{P}'', C'', \boldsymbol{Z}'', X'')$ *the laws of* X' *and* X'' *agree, that is to say* $X'[\mathbb{P}'] = X''[\mathbb{P}'']$.

Let us fix $f_\eta[\boldsymbol{Z}, X]_{.-} * Z^\eta$ to be the universal integral $f_\eta[\boldsymbol{Z}, X]_{.-} \circledast Z^\eta$ of remarks 3.7.27 and represent $(C_., X_., \boldsymbol{Z}_.)$ on canonical path space $\mathscr{D}^{n+d+n}$ in the manner of item 2.3.11. The image of $\mathbb{P}'$ under the representation is then a probability $\overline{\mathbb{P}}'$ on $\mathscr{D}^{n+d+n}$ that is carried by the "universal solution set"

$$S \stackrel{\text{def}}{=} \Big\{(c_., \boldsymbol{z}_., x_.) :\ x_. = c_. + \big(\boldsymbol{f}[\boldsymbol{z}_., x_.]_{.-} \circledast \boldsymbol{z}\big)_.\Big\} \tag{5.5.3}$$

and whose projection on the " $(C, \boldsymbol{Z})$-component" $\mathscr{D}^{n+d}$ is the law $\mathbb{L}$ of $(C, \boldsymbol{Z})$. Doing this to another weak solution Ξ'' will only change the measure from $\overline{\mathbb{P}}'$ to $\overline{\mathbb{P}}''$. The uniqueness problem turns into the question of whether the solution set S supports different probabilities whose projection on $\mathscr{D}^{n+d}$ is $\mathbb{L}$. Our equation will have a strong solution precisely if there is an *adapted* cross section $\mathscr{D}^{n+d} \to S$. We shall henceforth adopt this picture but write the evaluation processes as $\boldsymbol{Z}_t(c_., \boldsymbol{z}_., x_.) = \boldsymbol{z}_t$, etc., without overbars.

We shall show below that there exist weak solutions to (5.5.2) when $\boldsymbol{Z}$ is continuous and $\boldsymbol{f}$ is endogenous and continuous and has at most linear growth (see theorem 5.5.4 on page 333). This is accomplished by generalizing to the stochastic case the usual proof involving Peano's method of little straight steps. The uniqueness is rather more difficult to treat and has been established only in much more restricted circumstances – when the driver has the special form of condition 5.4.28 and the coupling coefficient is markovian and suitably nondegenerate; below we give two proofs (theorem 5.5.10 and exercise 5.5.14). For more we refer the reader to the literature ([105], [34], [54]).

The Size of the Solution

We continue to assume that $\boldsymbol{Z} = (Z^1, \ldots, Z^d)$ is a local L^q-integrator for some $q \geq 2$ and pick a $p \in [2, q]$. For a suitable choice of M (see (5.2.26)), the arguments of items 5.1.4 and 5.1.5 that led to the inequalities (5.1.16) and (5.1.17) on page 276 provide the a priori estimates (5.2.23) and (5.2.24) of the size of the solution X. They were established using the Lipschitz nature of the coupling coefficient $\boldsymbol{F}$ in an essential way. We shall now prove an a priori growth estimate that assumes no Lipschitz property, merely ***linear growth:*** there exist constants A, B such that up to evanescence

$$\left|\boldsymbol{F}[X]\right|_{\infty p} \leq A + B \cdot \left|X^\star\right|_p . \tag{5.5.4}$$

This implies $\qquad \left|\boldsymbol{F}[X]_T\right|_{\infty p} \leq A + B \cdot \left|X^\star_T\right|_p$

for all stopping times T, and in particular

$$\left|\boldsymbol{F}[X]_{T^\lambda -}\right|_{\infty p} \leq A + B \cdot \left|X^\star_{T^\lambda -}\right|_p$$

for the stopping times T^λ of THE time transformation, which in turn implies

$$\Big\|\left|\boldsymbol{F}[X]_{T^\lambda -}\right|_{\infty p}\Big\|_{L^p} \leq A + B \cdot \Big\|\left|X^\star_{T^\lambda -}\right|_p\Big\|_{L^p} \quad \forall\, \lambda > 0 . \tag{5.5.5}$$

This last is the form in which the assumption of linear growth enters the arguments. We will discuss this in the context of the general equation (5.2.18) on page 289:

$$X = C + F_\eta[X]_{\text{.}-} {*} Z^\eta . \tag{5.5.6}$$

Lemma 5.5.3 *Assume that X is a solution of (5.5.6), that the coupling coefficient $\boldsymbol{F}$ satisfies the linear-growth condition (5.5.5), and that*[43]

$$\Big\|\left|C^\star_{T^\lambda -}\right|_p\Big\|_{L^p} < \infty \quad \textit{and} \quad \Big\|\left|X^\star_{T^\lambda -}\right|_p\Big\|_{L^p} < \infty \tag{5.5.7}$$

for all $\lambda > 0$. Then there exists a constant $M = M_{p;B}$ such that

$$\|X\|^\star_{p,M} \leq 2\Big(A/B + \sup_{\lambda > 0} \Big\|\left|C^\star_{T^\lambda -}\right|_p\Big\|_{L^p}\Big) . \tag{5.5.8}$$

[43] If (5.5.4) holds, then inequality (5.5.7) can of course always be had provided we are willing to trade the given probability for a suitable equivalent one and to argue only up to some finite stopping time (see theorem 4.1.2).

Proof. Set $\Delta \stackrel{\text{def}}{=} X - C$ and let $0 \le \kappa < \mu$. Let S be a stopping time with $T^\kappa \le S < T^\mu$ on $[T^\kappa < T^\mu]$. Such S exist arbitrarily close to T^μ due to the predictability of that stopping time. Then

$$\Big\|\Big|(\Delta-\Delta^{T^\kappa})^\star_S\Big|_p\Big\|_{L^p} \le \Big\|\Big|(((T^\kappa, S]] \cdot \boldsymbol{F}[X]_{\boldsymbol{.}-}*\boldsymbol{Z})^\star_S\Big|_p\Big\|_{L^p} \le C_p^{\diamond(4.5.1)} \cdot \big|Q\big|_p ,$$

where
$$Q^\nu \stackrel{\text{def}}{=} \max_{\rho=1^\diamond,p^\diamond} \Big\|\Big(\int_{T^\kappa}^S \sup_\eta \big|F^\nu_\eta[X]\big|^\rho_{s-}\, d\Lambda_s\Big)^{1/\rho}\Big\|_{L^p}$$

$$\le \max_{\rho=1^\diamond,p^\diamond} \Big\|\Big(\int_\kappa^\mu \sup_\eta \big|F^\nu_\eta[X]\big|^\rho_{T^\lambda-}\, d\lambda\Big)^{1/\rho}\Big\|_{L^p} .$$

Thus
$$\big|Q\big|_p \le \max_{\rho=1^\diamond,p^\diamond} \Big\|\Big|\Big(\int_\kappa^\mu \sup_\eta \big|F_\eta[X]\big|^\rho_{T^\lambda-}\, d\lambda\Big)^{1/\rho}\Big|_p\Big\|_{L^p} ,$$

using A.3.29:
$$\le \max_{\rho=1^\diamond,p^\diamond} \Big(\int_\kappa^\mu \Big\|\Big|\sup_\eta \big|F_\eta[X]\big|_{T^\lambda-}\Big|_p\Big\|^\rho_{L^p}\, d\lambda\Big)^{1/\rho}$$

using (5.5.5):
$$\le \max_{\rho=1^\diamond,p^\diamond} \Big(\int_\kappa^\mu \big(A + B\big\|\big|X^\star_{T^\lambda-}\big|_p\big\|_{L^p}\big)^\rho\, d\lambda\Big)^{1/\rho} .$$

Taking S through a sequence announcing T^μ gives

$$\Big\|\Big|(\Delta-\Delta^{T^\kappa})^\star_{T^\mu-}\Big|_p\Big\|_{L^p} \le C_p^\diamond \max_{\rho=1^\diamond,p^\diamond} \Big(\int_\kappa^\mu\big(A+B\big\|\big|X^\star_{T^\lambda-}\big|_p\big\|_{L^p}\big)^\rho\, d\lambda\Big)^{1/\rho} . \tag{5.5.9}$$

For $\kappa = 0$, we have $T^\kappa = 0$, $X_0 = C_0$, and $\Delta_0 = 0$, so (5.5.9) implies

$$\big\|\big|X^\star_{T^\mu-}\big|_p\big\|_{L^p} \le \big\|\big|C^*_{T^\mu-}\big|_p\big\|_{L^p} + C_p^\diamond \max_{\rho=1^\diamond,p^\diamond} \Big(\int_0^\mu\big(A+B\big\|\big|X^\star_{T^\lambda-}\big|_p\big\|_{L^p}\big)^\rho\, d\lambda\Big)^{\frac{1}{\rho}} .$$

Gronwall's lemma in the form of exercise A.2.36 on page 384 now produces the desired inequality (5.5.8). ∎

Existence of Weak Solutions

Theorem 5.5.4 *Assume the driver $\boldsymbol{Z}$ is continuous; the coupling coefficient $\boldsymbol{f}$ is endogenous (p. 289) and non-anticipating, continuous,*[21] *and has at most linear growth; and the initial condition C is constant in time.*

Then the stochastic differential equation $X = C + \boldsymbol{f}[\boldsymbol{Z}, X]*\boldsymbol{Z}$ has a weak solution.

The proof requires several steps. The continuity of the driver entrains that THE previsible controller $\Lambda = \Lambda^{\langle q\rangle}[\boldsymbol{Z}]$ and the solution X of equation (5.5.6) are continuous as well. Both Λ and THE time transformation associated with it are now strictly increasing and continuous. Also, $X^\star_{T^\lambda-} = X^\star_{T^\lambda}$ for all λ, and $p^\diamond = 2$. Using inequality (5.5.8) and carrying out the λ-integral in (5.5.9) provides the inequality

$$\Big\|\Big|\big(X - X^{T^\kappa}\big)^\star_{T^\lambda}\Big|_p\Big\|_{L^p} \le c_\mu \cdot |\kappa - \lambda|^{1/2} , \qquad \kappa, \lambda \in [0, \mu] ,\ |\kappa - \lambda| < 1 ,$$

where $c_\mu = c_{\mu;A,B}$ is a constant that grows exponentially with μ and depends only on the indicated quantities $\mu; A, B$. We raise this to the p^{th} power and obtain

$$\mathbb{E}\Big[|X_{T^\kappa} - X_{T^\lambda}|_p^p\Big] \le c_\mu \cdot |\kappa - \lambda|^{p/2} . \tag{$*_1$}$$

The driver clearly satisfies a similar inequality:

$$\mathbb{E}\Big[|\boldsymbol{Z}_{T^\kappa} - \boldsymbol{Z}_{T^\lambda}|_p^p\Big] \le c'_\mu \cdot |\kappa - \lambda|^{p/2} . \tag{$*_2$}$$

We choose $p > 2$ and invoke Kolmogorov's lemma A.2.37 (ii) to establish as a first step toward the proof of theorem 5.5.4 the

Lemma 5.5.5 *Denote by $\mathbb{X}_{AB}$ the collection of all those solutions of equation (5.5.6) whose coupling coefficient satisfies inequality (5.5.5).*

(i) For every $\alpha < 1$ there exists a set C_α of paths in $\mathscr{C}^{d+n}$, compact[21] *and therefore uniformly equicontinuous on every bounded time-interval, such that*

$$\mathbb{P}\big[(\boldsymbol{Z}_., X_.) \in C_\alpha\big] > 1 - \alpha , \qquad\qquad X_. \in \mathbb{X}_{AB} .$$

(ii) Therefore the set $\big\{(\boldsymbol{Z}_., X_.)[\mathbb{P}] : X_. \in \mathbb{X}_{AB}\big\}$ of laws on $\mathscr{C}^{d+n}$ is uniformly tight and thus is relatively compact.[44]

Proof. Fix an instant u. There exists a $\mu > 0$ so that $\Omega^\Lambda \overset{\text{def}}{=} [\Lambda_u \le \mu] = [T^\mu \ge u]$ has $\mathbb{P}[\Omega^\Lambda] > 1 - \alpha/2$. As in the arguments of pages 14–15 we regard Λ as a $\mathbb{P}^*$-measurable map on $[\Lambda_u < \mu]$ whose codomain is $\mathscr{C}[0, u]$ equipped with the uniform topology. According to the definition 3.4.2 of measurability or Lusin's theorem, there exists a subset $\Omega^\Lambda_\alpha \subset \Omega^\Lambda$ with $\mathbb{P}[\Omega^\Lambda_\alpha] > 1 - \alpha/2$ on which Λ is uniformly continuous in the uniformity generated by the (idempotent) functions in $\mathcal{F}_u$, a uniformity whose completion is compact. Hence the collection $\Lambda_.(\Omega^\Lambda_\alpha)$ of increasing functions has compact closure $\underline{C}^\Lambda_\alpha$ in $\mathscr{C}[0, u]$.

For $X_. \in \mathbb{X}_{AB}$ consider the paths $\lambda \mapsto (\underline{\boldsymbol{Z}}_\lambda, \underline{X}_\lambda) \overset{\text{def}}{=} (\boldsymbol{Z}_{T^\lambda}, X_{T^\lambda})$ on $[0, \mu]$. Kolmogorov's lemma A.2.37 in conjunction with $(*_1)$ and $(*_2)$ provides a compact set $\underline{C}^{AB}_\alpha$ of continuous paths $\lambda \mapsto (\boldsymbol{z}_\lambda, x_\lambda)$, $0 \le \lambda \le \mu$, such that the set $\Omega^X_\alpha \overset{\text{def}}{=} \big[(\underline{\boldsymbol{Z}}^\mu_., \underline{X}^\mu_.) \in \underline{C}^{AB}_\alpha\big]$ has $\mathbb{P}[\Omega^X_\alpha] > 1 - \alpha/2$ simultaneously for every $X_.$ in $\mathbb{X}_{AB}$. Since the paths of $\underline{C}^{AB}_\alpha$ are uniformly equicontinuous (exercise A.2.38), the composition map $\circ$ on $\underline{C}^{AB}_\alpha \times \underline{C}^\Lambda_\alpha$, which sends $\big((\underline{\boldsymbol{z}}_. \underline{x}_.), \lambda_.\big)$ to $t \mapsto (\underline{x}_{\lambda_t}, \underline{\boldsymbol{z}}_{\lambda_t})$, is continuous and thus has compact image $C_\alpha \overset{\text{def}}{=} \underline{C}^{AB}_\alpha \circ \underline{C}^\Lambda_\alpha \subset \mathscr{C}^{n+d}[0, u]$. Indeed, let $\epsilon > 0$. There is a $\delta > 0$ so that $|\lambda' - \lambda| < \delta$ implies $|(\boldsymbol{z}_{\lambda'}, x_{\lambda'}) - (\boldsymbol{z}_\lambda), x_\lambda| < \epsilon/2$ for all $(\boldsymbol{z}_., x_.) \in \underline{C}^{AB}_\alpha$ and all $\lambda, \lambda' \in [0, \mu]$. If $|\lambda' - \lambda| < \delta$ in C^Λ_α and $|(\boldsymbol{z}'_., x'_.) - (\boldsymbol{z}_., x_.)| < \epsilon/2$,

[44] The pertinent topology on the space of probabilities on path spaces is the topology of weak convergence of measures; see section A.4.

then $|(z'_{\lambda'_t}, x'_{\lambda'_t}) - (z_{\lambda_t}, x_{\lambda_t})| \le |(z'_{\lambda'_t}, x'_{\lambda'_t}) - (z_{\lambda'_t}, x_{\lambda'_t})|\,|(z_{\lambda'_t}, x_{\lambda'_t}) - (z_{\lambda_t}, x_{\lambda_t})|$ $< \epsilon + \epsilon = 2\epsilon$; taking the supremum over $t \in [0,u]$ yields the claimed continuity. Now on $\Omega_\alpha \overset{\text{def}}{=} \Omega^\Lambda_\alpha \cap \Omega^X_\alpha$ we have clearly $(\underline{Z}_{\Lambda_t}, \underline{X}_{\Lambda_t}) = (Z_t, X_t)$, $0 \le t \le u$. That is to say, $(Z_., X_.)$ maps the set Ω_α, which has $\mathbb{P}[\Omega_\alpha] > 1 - \alpha$, into the compact set $C_\alpha \subset \mathscr{C}^{d+n}[0,u]$, a set that was manufactured from Z and A, B, p alone.

Since $\alpha < 1$ was arbitrary, the set of laws $\{(Z_., X_.)[\mathbb{P}] : X_. \in \mathbb{X}_{AB}\}$ is uniformly tight and thus (proposition A.4.6) is relatively compact.[44]

Actually, so far we have shown only that the projections on $\mathscr{C}^{d+n}[0,u]$ of these laws form a relatively weakly compact set, for any instant u. The fact that they form a relatively compact[44] set of probabilities on $\mathscr{C}^{d+n}[0,\infty)$ and are uniformly tight is left as an exercise. ∎

Proof of Theorem 5.5.4. For $n \in \mathbb{N}$ let $\mathcal{S}^{(n)}$ be the partition $\{k2^{-n} : k \in \mathbb{N}\}$ of time, define the coupling coefficient $F^{(n)}$ as the $\mathcal{S}^{(n)}$-scalæfication of f, and consider the corresponding stochastic differential equation

$$X^{(n)}_t = C + \int_0^t F^{(n)}_s[X^{(n)}]\, dZ_s = C + \sum_{0 \le k} f[Z, X^{(n)}]_{k2^{-n}} \cdot \big(Z_t - Z_{k2^{-n}\wedge t}\big) \,.$$

It has a unique solution, obtained recursively by $X^{(n)}_0 = C$ and

$$X^{(n)}_t = X^{(n)}_{k2^{-n}} + f[Z^{k2^{-n}}_., X^{(n)k2^{-n}}_.]_{k2^{-n}} \cdot \big(Z_t - Z_{k2^{-n}}\big) \qquad (5.5.10)$$
$$\text{for } \ k2^{-n} \le t \le (k{+}1)2^{-n} \,, \qquad k = 0, 1, \dots \,.$$

For later use note here that the map $Z_. \mapsto X^{(n)}_.$ is evidently continuous.[21] Also, the linear-growth assumption $|f[z,x]_t| \le A + B \cdot x^\star_t$ implies that the $F^{(n)}$ all satisfy the linear-growth condition (5.5.4). The laws $\mathbb{L}^{(n)}$ of the $(Z_., X^{(n)}_.)$ on path space $\mathcal{C}^{d+n}[0,\infty)$ form, in view of lemma 5.5.5, a relatively compact[44] set of probabilities. Extracting a subsequence and renaming it to $(\mathbb{L}^{(n)})$ we may assume that this sequence converges[44] to a probability $\mathbb{L}'$ on $\mathcal{C}^{n+d}[0,\infty)$. We set $\Omega' \overset{\text{def}}{=} \mathbb{R}^n \times \mathcal{C}^{d+n}[0,\infty)$ and $\mathbb{P}' \overset{\text{def}}{=} C[\mathbb{P}] \times \mathbb{L}'$ and equip Ω' with its canonical filtration $\mathcal{F}'$. On it there live the natural processes $Z'_., X'_.$ defined by

$$Z'_t\big((c, z_., x_.)\big) \overset{\text{def}}{=} z_t \,, \quad \text{and} \quad X'_t\big((c, z_., x_.)\big) \overset{\text{def}}{=} x_t \qquad t \ge 0 \,,$$

and the random variable $C' : (c, z_., x_.) \mapsto c$. If we can show that, under $\mathbb{P}'$, $X' = C + f[Z', X']{*}Z'$, then the theorem will be proved; that $(C', Z'_.)$ has the same distribution under $\mathbb{P}'$ as $(C, Z_.)$ has under $\mathbb{P}$, that much is plain.

Let us denote by $\mathbb{E}'$ and $\mathbb{E}^{(n)}$ the expectations with respect to $\mathbb{P}'$ and $\mathbb{P}^{(n)} \overset{\text{def}}{=} C[\mathbb{P}] \times \mathbb{L}^{(n)}$, respectively. Below we will need to know that Z' is a $\mathbb{P}'$-integrator:

$$\lceil\!\lceil Z' \rceil\!\rceil_{\mathcal{I}^p[\mathbb{P}']} \le \lceil\!\lceil Z \rceil\!\rceil_{\mathcal{I}^p[\mathbb{P}]} \,. \qquad (5.5.11)$$

To see this let $\mathcal{A}_t$ denote the algebra of bounded continuous functions $f : \Omega' \to \mathbb{R}$ that depend only on the values of the path at finitely many instants s prior to t; such f is the composition of a continuous bounded function ϕ on $\mathbb{R}^{(n+d+n)\times k}$ with a vector $(c, \boldsymbol{z}_{\boldsymbol{\cdot}}, x_{\boldsymbol{\cdot}}) \mapsto \big((c, \boldsymbol{z}_{s_i}, x_{s_i}) : 1 \le i \le k\big)$. Evidently $\mathcal{A}_t$ is an algebra and vector lattice closed under chopping that generates $\mathcal{F}_t$. To see (5.5.11) consider an elementary integrand $\boldsymbol{X}'$ on $\mathcal{F}'$ whose d components X'_η are as in equation (2.1.1) on page 46, but special in the sense that the random variables $X'_{\eta s}$ belong to $\mathcal{A}_s$, at all instants s. Consider only such $\boldsymbol{X}'$ that vanish after time t and are bounded in absolute value by 1. An inspection of equation (2.2.2) on page 56 shows that, for such $\boldsymbol{X}'$, $\int \boldsymbol{X}' \, d\boldsymbol{Z}'$ is a continuous function on Ω'. The composition of $\boldsymbol{X}'$ with $(\boldsymbol{Z}_{\boldsymbol{\cdot}}, X_{\boldsymbol{\cdot}})$ is a previsible process $\boldsymbol{X}$ on $(\Omega, \mathcal{F})$ with $|\boldsymbol{X}| \le [\![0, t]\!]$, and

$$\begin{aligned}\mathbb{E}'\Big[\Big|\int \boldsymbol{X}' \, d\boldsymbol{Z}'\Big|^p \wedge K\Big] &= \lim \mathbb{E}^{(n)}\Big[\Big|\int \boldsymbol{X}' \, d\boldsymbol{Z}'\Big|^p \wedge K\Big] \\ &= \lim \mathbb{E}\Big[\Big|\int \boldsymbol{X}^{(n)} \, d\boldsymbol{Z}\Big|^p \wedge K\Big] \le \big\lceil\!\big\lceil \boldsymbol{Z}^t \big\rceil\!\big\rceil^p_{\mathcal{I}^p[\mathbb{P}]} . \qquad (5.5.12)\end{aligned}$$

We take the supremum over $K \in \mathbb{N}$ and apply exercise 3.3.3 on page 109 to obtain inequality (5.5.11).

Next let $t \ge 0$, $\alpha \in (0,1)$, and $\epsilon > 0$ be given. There exists a compact[21] subset $C_\alpha \in \mathcal{F}'_t$ such that $\mathbb{P}'[C_\alpha] > 1 - \alpha$ and $\mathbb{P}^{(n)}[C_\alpha] > 1 - \alpha \quad \forall\, n \in \mathbb{N}$. Then

$$\begin{aligned}&\mathbb{E}'\Big[\big|X' - \big(C' + \boldsymbol{f}[\boldsymbol{Z}', X'] {*} \boldsymbol{Z}'\big)\big|^\star_t \wedge 1\Big] \\ &\le \quad \mathbb{E}'\Big[\Big|\big(\boldsymbol{f}[\boldsymbol{Z}', X'] - \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X']\big) {*} \boldsymbol{Z}'\Big|^\star_t \wedge 1\Big] \\ &\qquad + \mathbb{E}'\Big[\Big|X' - \Big(C' + \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X'] {*} \boldsymbol{Z}'\Big)\Big|^\star_t \wedge 1\Big] \\ &\le \quad \mathbb{E}'\Big[\Big|\big(\boldsymbol{f}[\boldsymbol{Z}', X'] - \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X']\big) {*} \boldsymbol{Z}'\Big|^\star_t \wedge 1\Big] \\ &\qquad + \big(\mathbb{E}' - \mathbb{E}^{(m)}\big)\Big[\Big|X' - \Big(C' + \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X'] {*} \boldsymbol{Z}'\Big)\Big|^\star_t \wedge 1\Big] \\ &\qquad + \mathbb{E}^{(m)}\Big[\Big|X' - \Big(C' + \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X'] {*} \boldsymbol{Z}'\Big)\Big|^\star_t \wedge 1\Big]\end{aligned}$$

Since $\mathbb{E}^{(m)}\big[\,\big|\,X' - \big(C' + \boldsymbol{f}^{(m)}[X', \boldsymbol{Z}'] {*} \boldsymbol{Z}'\big)\,\big|^\star_t \wedge 1\big] = 0$:

$$\begin{aligned}&\le \quad \mathbb{E}'\Big[\Big|\big(\boldsymbol{f}[\boldsymbol{Z}', X'] - \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X']\big) {*} \boldsymbol{Z}'\Big|^\star_t \wedge 1\Big] \\ &\qquad + \big(\mathbb{E}' - \mathbb{E}^{(m)}\big)\Big[\Big|X' - \Big(C' + \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X'] {*} \boldsymbol{Z}'\Big)\Big|^\star_t \wedge 1\Big] \\ &\qquad + \mathbb{E}^{(m)}\Big[\Big|\big(\boldsymbol{f}^{(m)}[X', \boldsymbol{Z}'] - \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X']\big) {*} \boldsymbol{Z}'\Big|^\star_t \wedge 1\Big] \\ &\le \quad 2\alpha + \mathbb{E}'\Big[\Big|\big(\boldsymbol{f}[\boldsymbol{Z}', X'] - \boldsymbol{f}^{(n)}[\boldsymbol{Z}', X']\big) {*} \boldsymbol{Z}'\Big|^\star_t \cdot C_\alpha\Big] \qquad (5.5.13)\end{aligned}$$

$$+\left(\mathbb{E}-\mathbb{E}^{(m)}\right)\left[\left|X'-\left(C'+\boldsymbol{f}^{(n)}[\boldsymbol{Z}',X']{*}\boldsymbol{Z}'\right)\right|_t^\star\wedge 1\right] \tag{5.5.14}$$

$$+\,\mathbb{E}^{(m)}\left[\left|\left(\boldsymbol{f}^{(m)}[X',\boldsymbol{Z}']-\boldsymbol{f}^{(n)}[\boldsymbol{Z}',X']\right){*}\boldsymbol{Z}'\right|_t^\star\cdot C_\alpha\right]. \tag{5.5.15}$$

Now the image under $\boldsymbol{f}$ of the compact set C_α is compact, on acount of the stipulated continuity of $\boldsymbol{f}$, and thus is uniformly equicontinuous (exercise A.2.38). There is an index N such that $|\,\boldsymbol{f}\big((x_{\boldsymbol{.}},\boldsymbol{z}_{\boldsymbol{.}})\big)-\boldsymbol{f}^{(n)}\big((x_{\boldsymbol{.}},\boldsymbol{z}_{\boldsymbol{.}})\big)\,|\le\epsilon$ for all $n\ge N$ and all $(x_{\boldsymbol{.}},\boldsymbol{z}_{\boldsymbol{.}})\in C_\alpha$. Since $\boldsymbol{f}$ is non-anticipating, $\boldsymbol{f}_{\boldsymbol{.}}[X_{\boldsymbol{.}},\boldsymbol{Z}_{\boldsymbol{.}}]$ is a continuous adapted process and so is predictable. So is $\boldsymbol{f}^{(n)}_{\boldsymbol{.}}[X_{\boldsymbol{.}},\boldsymbol{Z}_{\boldsymbol{.}}]$. Therefore $|\,\boldsymbol{f}-\boldsymbol{f}^{(n)}\,|\le\epsilon$ on the predictable envelope $\widehat{C}_\alpha$ of C_α. We conclude with exercise 3.7.16 on page 137 that $\big(\boldsymbol{f}[\boldsymbol{Z}',X']-\boldsymbol{f}^{(n)}[\boldsymbol{Z}',X']\big){*}\boldsymbol{Z}'$ and $\big((\boldsymbol{f}[\boldsymbol{Z}',X']-\boldsymbol{f}^{(n)}[\boldsymbol{Z}',X'])\cdot\widehat{C}_\alpha\big){*}\boldsymbol{Z}'$ agree on C_α. Now the integrand of the previous indefinite integral is uniformly less than ϵ, so the maximal inequality (2.3.5) furnishes the inequality

$$\mathbb{E}'\left[\left|\left(\boldsymbol{f}[\boldsymbol{Z}',X']-\boldsymbol{f}^{(n)}[\boldsymbol{Z}',X']\right){*}\boldsymbol{Z}'\right|_t^\star\cdot C_\alpha\right]\le\epsilon\cdot C_1^\star\,\lceil\!\lceil \boldsymbol{Z}'^t\rceil\!\rceil_{\mathcal{I}^1[\mathbb{P}']}$$

by inequality (5.5.11): $\qquad\le\epsilon\cdot C_1^\star\,\lceil\!\lceil \boldsymbol{Z}^t\rceil\!\rceil_{\mathcal{I}^1[\mathbb{P}]}\,.$

The term in (5.5.15) can be estimated similarly, so that we arrive at

$$\begin{aligned}\mathbb{E}'\left[\left|X'-\left(C'+\boldsymbol{f}[\boldsymbol{Z}',X']{*}\boldsymbol{Z}'\right)_t^\star\right|\wedge 1\right]&\le 2\alpha+2\epsilon\cdot C_1^\star\,\lceil\!\lceil \boldsymbol{Z}^t\rceil\!\rceil_{\mathcal{I}^1[\mathbb{P}]}\\&\quad+\left(\mathbb{E}-\mathbb{E}^{(m)}\right)\left[\left|X'-\left(C'+\boldsymbol{f}^{(n)}[\boldsymbol{Z}',X']{*}\boldsymbol{Z}'\right)\right|_t^\star\wedge 1\right].\end{aligned}$$

Now the expression inside the brackets $[\;]$ of the previous line is a continuous bounded function on $\mathscr{C}^{d+n}$ (see equation (5.5.10)); by the choice of a sufficiently large $m\ge N$ it can be made arbitrarily small. In view of the arbitrariness of α and ϵ, this boils down to $\mathbb{E}'\Big[\big|X'-\big(C'+\boldsymbol{f}[\boldsymbol{Z}',X']{*}\boldsymbol{Z}'\big)\big|_t^\star\wedge 1\Big]=0$. ∎

Problem 5.5.6 *Find a generalization to non–continuous drivers* $\boldsymbol{Z}$.

Uniqueness

The known uniqueness results for weak solutions cover mainly what might be called the "Classical Stochastic Differential Equation"

$$X_t=x+\int_0^t f_0(s,X_s)\,ds+\sum_{\eta=1}^d\int_0^t f_\eta(s,X_s)\,dW_s^\eta\,. \tag{5.5.16}$$

Here the driver is as in condition 5.4.28. They all require the ***uniform ellipticity***[7] of the symmetric matrix

$$a^{\mu\nu}(t,x)\stackrel{\text{def}}{=}\frac12\sum_{\eta=1}^d f_\eta^\mu(t,x)f_\eta^\nu(t,x)\,,$$

namely, $\qquad a^{\mu\nu}(t,x)\xi_\mu\xi_\nu\ge\beta^2\cdot|\xi|^2\qquad\forall\,\xi,x\in\mathbb{R}^n\,,\quad\forall\,t\in\mathbb{R}_+ \qquad (5.5.17)$

for some $\beta > 0$. We refer the reader to [105] for the most general results. Here we will deal only with the "Classical Time-Homogeneous Stochastic Differential Equation"

$$X_t = x + \int_0^t f_0(X_s)\, ds + \sum_{\eta=1}^{d} \int_0^t f_\eta(X_s)\, dW_s^\eta \tag{5.5.18}$$

under stronger than necessary assumptions on the coefficients f_η. We give two uniqueness proofs, mostly to exhibit the connection that stochastic differential equations (SDEs) have with elliptic and parabolic partial differential equations (PDEs) of order 2.

The uniform ellipticity can of course be had only if the dimension d of $\boldsymbol{W}$ exceeds the dimension n of the state space; it is really no loss of generality to assume that $n = d$. Then the matrix $f_\eta^\nu(x)$ is invertible at all $x \in \mathbb{R}^n$, with a uniformly bounded inverse named $F(x) \stackrel{\text{def}}{=} f^{-1}(x)$.

We shall also assume that the f_η^ν are continuous and bounded. For ease of thinking let us use the canonical representation of page 331 to shift the whole situation to the path space $\mathscr{C}^n$. Accordingly, the value of X_t at a path $\omega = x_{\boldsymbol{.}} \in \mathscr{C}^n$ is x_t. Because of the identity

$$W_t^\eta = \textstyle\sum_\nu \displaystyle\int_0^t F_\nu^\eta(X_s)\ \left(dX_s^\nu - f_0^\nu(X_s)ds\right)\,, \tag{5.5.19}$$

$\boldsymbol{W}$ is adapted to the natural filtration on path space. In this situation the problem becomes this: denoting by $\mathfrak{P}$ the collection of all probabilities on $\mathscr{C}^n$ under which the process $\boldsymbol{W}_t$ of (5.5.19) is a standard Wiener process, show that $\mathfrak{P}$ – which we know from theorem 5.5.4 to be non-void – is in fact a singleton.

There is no loss of generality in assuming that $f_0 = 0$, so that equation (5.5.18) turns into

$$X_t = x + \sum_{\eta=1}^{d} \int_0^t f_\eta(X_s)\, dW_s^\eta\ . \tag{5.5.20}$$

Exercise 5.5.7 Indeed, one can use Girsanov's theorem 3.9.19 to show the following: If the law of any process X satisfying (5.5.20) is unique, then so is the law of any process X satisfying (5.5.18).

All of the known uniqueness proofs also have in common the need for some input in the form of hard estimates from Fourier analysis or PDE. The proof given below may have some little whimsical appeal in that it does not refer to the martingale problem ([105], [34], and [54]) but uses the existence of solutions of the Dirichlet problem for its outside input. A second slightly simpler proof is outlined in exercises 5.5.13–5.5.14.

5.5.8 The Dirichlet Problem in its form pertinent to the problem at hand is to find for a given domain B of $\mathbb{R}^n$ a *continuous* function $u : \overline{B} \to \mathbb{R}$ with two *continuous* derivatives in the interior $\mathring{B}$ that solves the PDE[40]

$$\mathcal{A}u(x) \stackrel{\text{def}}{=} \frac{a^{\mu\nu}(x)}{2}\, u_{;\mu\nu}(x) = 0 \quad \forall\, x \in \mathring{B} \tag{5.5.21}$$

and satisfies the boundary condition

$$u(x) = g(x) \quad \forall\, x \in \partial B \stackrel{\text{def}}{=} \overline{B} \backslash \mathring{B}\ . \tag{5.5.22}$$

If a is the identity matrix, then this is the classical Dirichlet problem asking for a function u harmonic inside B, continuous on $\overline{B}$, and taking the prescribed value g on the boundary. This problem has a unique solution if B is a box and g is continuous; it can be constructed with the time-honored method of separation of variables, which the reader has seen in third-semester calculus. The solution of the classical problem can be parlayed into a solution of (5.5.21)–(5.5.22) when the coefficient matrix $a(x)$ is continuous, the domain B is a box, and the boundary value g is smooth ([37]). For the sake of accountability we put this result as an assumption on $\boldsymbol{f}$:

Assumption 5.5.9 *The coefficients $f^\mu_\eta(x)$ are continuous and bounded and (i) the matrix $a^{\mu\nu}(x) \stackrel{\text{def}}{=} \sum_\eta f^\mu_\eta(x) f^\nu_\eta(x)$ satisfies the strict ellipticity (5.5.17); (ii) the Dirichlet problem (5.5.21)–(5.5.22) with smooth boundary data g has a solution of class $C^2(\mathring{B}) \cap C^0(\overline{B})$ on every box B in $\mathbb{R}^n$ whose sides are perpendicular to the axes.*

The connection of our uniqueness quest with the Dirichlet problem is made through the following observation. Suppose that X^x is a weak solution of the stochastic differential equation (5.5.20). Let u be a solution of the Dirichlet problem above, with B being some relatively compact domain in $\mathbb{R}^n$ containing the point x in its interior. By exercise 3.9.10, the first time T at which X^x hits the boundary of the domain is almost surely finite, and Itô's formula gives

$$\begin{aligned} u(X^x_T) &= u(X^x_0) + \int_0^T u_{;\nu}(X^x_s)\, dX^{x\nu}_s + \frac{1}{2}\int_0^T u_{;\mu\nu}(X^x_s)\, d[X^{x\mu}, X^{x,\nu}]_s \\ &= u(x) + \int_0^T u_{;\nu}(X^x_s) f^\nu_\eta(X^x_s)\, dW^\eta\ , \end{aligned}$$

since $\mathcal{A}u = 0$ and thus $u_{;\mu\nu}(X^x_s) d[X^{x\mu}, X^{x\nu}]_s = 2\mathcal{A}u(X^x_s)ds = 0$ on $[s < T]$. Now u, being continuous on $\overline{B}$, is bounded there. This exhibits the right-hand side as the value at T of a bounded local martingale. Finally, since $u = g$ on ∂B,

$$u(x) = \mathbb{E}\big[g(X^x_T)\big]\ . \tag{5.5.23}$$

This equality provides two uniqueness statements: the solution u of the Dirichlet problem, for whose existence we relied on the literature, is unique;

indeed, the equality expresses $u(x)$ as a construct of the vector fields f_η and the boundary function g. We can also read off the ***maximum principle***: u takes its maximum and its minimum on the boundary ∂B. The uniqueness of the solution implies at the same time that the map $g \mapsto u(x)$ is linear. Since it satisfies $|u(x)| \le \sup\{|g(x)| : x \in \partial B\}$ on the algebra of functions g that are restrictions to ∂B of smooth functions, an algebra that is uniformly dense in $C(\partial B)$ by theorem A.2.2 (iii), it has a unique extension to a continuous linear map on $C(\partial B)$, a Radon measure. This is called the ***harmonic measure*** for the problem (5.5.21)–(5.5.22) and is denoted by $\eta^x_{\partial B}(d\sigma)$.

The second uniqueness result concerns any probability $\mathbb{P}$ under which X is a weak solution of equation (5.5.20). Namely, (5.5.23) also says that the ***hitting distribution*** of X^x on the boundary ∂B, by which we mean the law of the process X^x at the first time T it hits the boundary, or the distribution $\lambda^x_{\partial B}(d\sigma) \overset{\text{def}}{=} X^x_T[\mathbb{P}]$ of the ∂B-valued random variable X^x_T, is determined by the matrix $a^{\mu\nu}(x)$ alone. In fact it is harmonic measure:

$$\lambda^x_{\partial B} \overset{\text{def}}{=} X^x_T[\mathbb{P}] = \eta^x_{\partial B} \qquad\qquad \forall\, \mathbb{P} \in \mathfrak{P} .$$

Look at things this way: varying B but so that $x \in \mathring{B}$ will produce lots of hitting distributions $\lambda^x_{\partial B}$ that are all images of $\mathbb{P}$ under various maps X^x_T but do actually not depend on $\mathbb{P}$. Any other $\mathbb{P}'$ under which X^x solves equation (5.5.20) will give rise to exactly the same hitting distributions $\lambda^x_{\partial B}$. Our goal is to parlay this observation into the uniqueness $\mathbb{P} = \mathbb{P}'$:

Theorem 5.5.10 *Under assumption 5.5.9, equation (5.5.18) has a unique weak solution.*

Proof. Only the uniqueness is left to be established. Let $H^\nu_{\ell,k}$ be the hyperplane in $\mathbb{R}^n$ with equation $x^\nu = k2^{-n}$, $\nu = 1, \ldots, n$, $1 \le \ell \in \mathbb{N}$, $k \in \mathbb{Z}$. According to exercise 3.9.10 on page 162, we may remove from $\mathscr{C}^n$ a $\mathfrak{P}$-nearly empty set N such that on the remainder $\mathscr{C}^n \backslash N$ the stopping times $S^\nu_{\ell,k} \overset{\text{def}}{=} \inf\{t : X_t \in H^\nu_{\ell,k}\}$ are continuous.[21] The random variables $X_{S^\nu_{\ell,k}}$ will then be continuous as well. Then we may remove a further $\mathfrak{P}$-nearly empty set N' such that the stopping times

$$S^{\nu,\nu'}_{\ell,\ell',k,k'} \overset{\text{def}}{=} \inf\{t > S^\nu_{\ell,k} : X_t \in H^{\nu'}_{\ell',k'}\} ,$$

too, are continuous on $\Omega \overset{\text{def}}{=} \mathscr{C}^n \setminus (N \cup N')$, and so on. With this in place let us define for every $\ell \in \mathbb{N}$ the stopping times $T^\ell_0 = 0$,

$$T^\ell_\nu \overset{\text{def}}{=} \inf\{t > T^{\ell,\nu} : X_t \in \bigcup\nolimits_k H^\nu_{\ell,k}\}$$

and
$$T^\ell_{k+1} \overset{\text{def}}{=} \inf\{T^{\ell,\nu} : T^{\ell,\nu} > T^\ell_k\} , \qquad\qquad k = 0, 1, \ldots .$$

T^ℓ_{k+1} is the first time after T^ℓ_k that the path leaves the smallest box with sides in the $H^\nu_{\ell,k}$ that contains $X_{T^\ell_k}$ in its interior. The T^ℓ_k are continuous on Ω, and so are the maps $\omega \mapsto X_{T_k^\ell}(\omega)$. In this way we obtain for every $\ell \in \mathbb{N}$ and $\omega = x_. \in \Omega$ a discrete path $\boldsymbol{x}_.^{(\ell)} : \mathbb{N} \ni k \mapsto X_{T_k^\ell}(\omega)$ in $\ell^0_{\mathbb{R}^n}$. The map $\omega \to \boldsymbol{x}_.^{(\ell)}$ is clearly continuous from Ω to $\ell^0_{\mathbb{R}^n}$. Let us identify $\boldsymbol{x}_.^{(\ell)}$ with the path $x'_. \in \mathscr{C}^n$ that at time T_k^ℓ has the value $\boldsymbol{x}_k^{(\ell)} = X_{T_k^\ell}$ and is linear between T_k^ℓ and T_{k+1}^ℓ. The map $\boldsymbol{x}_.^{(\ell)} \mapsto x'_.$ is evidently continuous from $\ell^0_{\mathbb{R}^n}$ to $\mathscr{C}^n$. We leave it to the reader to check that for $\jmath \le \ell$ the times $T_k^\jmath$ agree on $x_.$ and $x'_.$, and that therefore $x_{T_k^\jmath} = x'_{T_k^\jmath}$, $\jmath \le \ell$, $1 \le k < \infty$. The point: for $\jmath \le \ell$

$$\boldsymbol{x}_.^{(\jmath)} \in \ell^0_{\mathbb{R}^n} \quad \text{is a continuous function of} \quad \boldsymbol{x}_.^{(\ell)} \in \ell^0_{\mathbb{R}^n} . \tag{5.5.24}$$

Next let $\mathcal{A}^{(\ell)}$ denote the algebra of functions on Ω of the form $x_. \mapsto \phi(\boldsymbol{x}_.^{(\ell)})$, where $\phi : \ell^0_{\mathbb{R}^n} \to \mathbb{R}$ is bounded and continuous. Equation (5.5.24) shows that $\mathcal{A}^{(\jmath)} \subset \mathcal{A}^{(\ell)}$ for $\jmath \le \ell$. Therefore $\mathcal{A} \stackrel{\text{def}}{=} \bigcup_\ell \mathcal{A}^{(\ell)}$ is an algebra of bounded continuous functions on Ω.

Lemma 5.5.11 *(i) If $x_.$ and $x'_.$ are two paths in Ω on which every function of $\mathcal{A}$ agrees, then $x_.$ and $x'_.$ describe the same arc (see definition 3.8.17). (ii) In fact, after removal from Ω of another $\mathfrak{P}$-nearly empty set $\mathcal{A}$ separates the points of Ω.*

Proof. (i) First observe that $x_0 = x'_0$. Otherwise there would exist a continuous bounded function ϕ on $\mathbb{R}^n$ that separates these two points. The function $x_. \mapsto \phi\big(\boldsymbol{x}_{T_0^\ell}\big)$ of $\mathcal{A}$ would take different values on $x_.$ and on $x'_.$. An induction in k using the same argument shows that $x_{T_k^\ell} = x'_{T_k^\ell}$ for all $\ell, k \in \mathbb{N}$. Given a $t > 0$ we now set

$$t' \stackrel{\text{def}}{=} \sup\{T_k^\ell(x'_.) : T_k^\ell(x_.) \le t\} .$$

Clearly $x_.$ and $x'_.$ describe the same arc via $t \mapsto t'$.

(ii) Using exercise 3.8.18 we adjust Ω so that whenever ω and ω' describe the same arc via $t \mapsto t'$ then, in view of equation (5.5.19), $\boldsymbol{W}_.(\omega)$ and $\boldsymbol{W}_.(\omega')$ also describe the same arc via $t \mapsto t'$, which forces $t = t'$ $\forall t$: any two paths of $X_.$ on which all the functions of $\mathcal{A}$ agree not only describe the same arc, they are actually identical. It is at this point that the differential equation (5.5.18) is used, through its consequence (5.5.19). ∎

Since every probability on the polish space $\mathscr{C}^n$ is tight, the uniqueness claim is immediate from proposition A.3.12 on page 399 once the following is established:

Lemma 5.5.12 *Any two probabilities in $\mathfrak{P}$ agree on $\mathcal{A}$.*

Proof. Let $\mathbb{P}, \mathbb{P}' \in \mathfrak{P}$, with corresponding expectations $\mathbb{E}, \mathbb{E}'$. We shall prove by induction in k the following: $\mathbb{E}$ and $\mathbb{E}'$ agree on functions in $\mathcal{A}^\ell$ of the form

$$\phi_0\big(X_{T_0^\ell}\big) \cdots \phi_k\big(X_{T_k^\ell}\big) , \tag{$*$}$$

$\phi_\kappa \in C_b(\mathbb{R}^n)$. This is clear if $k = 0$: $\phi_0\big(X_{T_0^\ell}\big) = \phi_0(x)$. We preface the induction step with a remark: $X_{T_k^\ell}$ is contained in a finite number of $n{-}1$-dimensional "squares" S^i of side length $2^{-\ell}$. About each of these there is a minimal box B^i containing S^i in its interior, and $X_{T_{k+1}^\ell}$ will lie in the union $\bigcup_i \partial B^i$ of their boundaries. Let u_{k+1}^i denote the solution of equation (5.5.21) on B^i that equals ϕ_{k+1} on ∂B^i. Then[35]

$$\begin{aligned}\phi_{k+1}\big(X_{T_{k+1}^\ell}\big)\cdot S^i{\circ}X_{T_k^\ell} &= u_{k+1}^i\big(X_{T_{k+1}^\ell}\big)\cdot S^i{\circ}X_{T_k^\ell}\\ &= u_{k+1}^i\big(X_{T_k^\ell}\big)\cdot S^i\circ X_{T_k^\ell} + \int_{T_k^\ell}^{T_{k+1}^\ell} u_{k+1;\nu}^i(X)\,dX^\nu\end{aligned}$$

has the conditional expectation

$$\mathbb{E}\Big[\phi_{k+1}\big(X_{T_{k+1}^\ell}\big)\cdot S^i{\circ}X_{T_k^\ell}\big|\mathcal{F}_{T_k^\ell}\Big] = u_{k+1}^i\big(X_{T_k^\ell}\big)\cdot S^i{\circ}X_{T_k^\ell}\ ,$$

whence

$$\mathbb{E}\Big[\phi_{k+1}\big(X_{T_{k+1}^\ell}\big)\big|\mathcal{F}_{T_k^\ell}\Big] = \textstyle\sum_i u_{k+1}^i\big(X_{T_k^\ell}\big)\ .$$

Therefore, after conditioning on $\mathcal{F}_{T_k^\ell}$,

$$\mathbb{E}\Big[\phi_0\big(X_{T_0^\ell}\big)\cdots\phi_{k+1}\big(X_{T_{k+1}^\ell}\big)\Big] = \mathbb{E}\Big[\phi_0\big(X_{T_0^\ell}\big)\cdots\Big(\phi_k\textstyle\sum_i u_{k+1}^i\Big)\big(X_{T_k^\ell}\big)\Big]\ .$$

By the same token

$$\mathbb{E}'\Big[\phi_0\big(X_{T_0^\ell}\big)\cdots\phi_{k+1}\big(X_{T_{k+1}^\ell}\big)\Big] = \mathbb{E}'\Big[\phi_0\big(X_{T_0^\ell}\big)\cdots\Big(\phi_k\textstyle\sum_i u_{k+1}^i\Big)\big(X_{T_k^\ell}\big)\Big]\ .$$

By the induction hypothesis the two right-hand sides agree. The induction is complete. Since the functions of the form $(*)$, $k \in \mathbb{N}$, form a multiplicative class generating $\mathcal{A}$, $\mathbb{E} = \mathbb{E}'$ on $\mathcal{A}$. The proof of the lemma is complete, and with it that of theorem 5.5.10. ∎

The next two exercise comprise another proof of the uniqueness theorem.

Exercise 5.5.13 The ***initial value problem*** for the differential operator $\mathcal{A}$ is the problem of finding, for every $\phi \in C_0(\mathbb{R}^n)$, a function $u(t,x)$ that is twice continuously differentiable in x and bounded on every strip $(0,t')\times\mathbb{R}^n$ and satisfies the ***evolution equation*** $\dot u = \mathcal{A}u$ ($\dot u$ denotes the t-partial $\partial u/\partial t$) and the initial condition $u(0,x) = \phi(x)$. Suppose X^x solves equation (5.5.20) under $\mathbb{P}$ and u solves the initial value problem. Then $[0,t'] \ni t \mapsto u(t'-t, X_t)$ is a martingale under $\mathbb{P}$.

Exercise 5.5.14 Retain assumption 5.5.9 (i) and assume that the initial value problem of exercise 5.5.13 has a solution in C^2 for every $\phi \in C_b^\infty(\mathbb{R}^n)$ (this holds if the coefficient matrix a is Hölder continuous, for example). Then again equation (5.5.18) has a unique weak solution.

5.6 Stochastic Flows

Consider now the situation that $\boldsymbol{Z}$ is an L^0-integrator, the coupling coefficient $\boldsymbol{F}$ is strongly Lipschitz, and that the initial condition a (constant) point $x \in \mathbb{R}^n$. We want to investigate how the solution X^x of

$$X_t^x = x + \int_0^t \boldsymbol{F}[X^x]_{s-}\, d\boldsymbol{Z}_s \tag{5.6.1}$$

depends on x. Considering x as the parameter in $U \stackrel{\text{def}}{=} \mathbb{R}^n$ and applying theorem 5.2.24, we may assume that $x \mapsto X_{\bullet}^x(\omega)$ is continuous from $\mathbb{R}^n$ to $\mathscr{D}^n$ equipped with the topology of uniform convergence on bounded intervals, for every $\omega \in \Omega$. In particular, the maps $\Xi_t = \Xi_t^\omega : x \mapsto X_t^x(\omega)$, one for every $\omega \in \Omega$ and every $t \geq 0$, map $\mathbb{R}^n$ continuously into itself. They constitute the ***stochastic flow*** that comes with (5.6.1). We want to investigate under which conditions the map Ξ_t^ω is a homeomorphism from $\mathbb{R}^n$ onto itself.

Assumption L *There exist constants L_η such that*

$$\left|F_\eta[X] - F_\eta[Y]\right| \leq L_\eta |X - Y| \,, \qquad 1 \leq \eta \leq d \,,$$

for any two adapted càdlàg $\mathbb{R}^n$-valued processes X, Y.

This is but the strong Lipschitz condition: inequality (5.2.7) is satisfied with $L^{(5.2.7)} \leq L \stackrel{\text{def}}{=} \sum_\eta L_\eta \leq nL^{(5.2.7)}$. It implies

$$\left|F_\eta[X] - F_\eta[Y]\right|_{\bullet-} \leq L_\eta |X - Y|_{\bullet-}^\star \,, \qquad 1 \leq \eta \leq d \,.$$

Here and in the remainder of this section $|\ |$ denotes the euclidean norm on $\mathbb{R}^n$ and $\langle\ |\ \rangle$ the inner product.

Assumption J *If $X^{(1)}, X^{(2)}$ are any two $\mathbb{R}^n$-valued solutions of $dX^{(i)} = x^{(i)} + F_\eta[X^{(i)}]_{\bullet-} {*} Z^\eta$ then the subset*

$$\begin{aligned}&\left[X_{\bullet-}^{(1)} \neq X_{\bullet-}^{(2)}\right] \cap \left[X^{(1)} = X^{(2)}\right] \\ = &\left[X_{\bullet-}^{(1)} \neq X_{\bullet-}^{(2)}\right] \cap \left[X_{\bullet-}^{(1)} + F_\eta[X^{(1)}]_{\bullet-} \Delta Z^\eta = X_{\bullet-}^{(2)} + F_\eta[X^{(2)}]_{\bullet-} \Delta Z^\eta\right]\end{aligned}$$

of the base space $\boldsymbol{B}$ is evanescent.

To paraphrase: for nearly every $\omega \in \Omega$, if at any instant s $X_{s-}^{(1)}(\omega)$ differs from $X_{s-}^{(2)}(\omega)$, then the effective jumps $F_\eta[X^{(1)}]_{s-}(\omega)\Delta Z_S^\eta(\omega)$ and $F_\eta[X^{(2)}]_{s-}(\omega)\Delta Z_S^\eta(\omega)$ will not propel both processes to the same point of $\mathbb{R}^n$. This assumption is clearly necessary if $\Xi_t^\omega : x \mapsto X_t^x(\omega)$ is to be injective for nearly all $\omega \in \Omega$.

Exercise Assumption J above is satisfied if $\boldsymbol{Z}$ ***has small jumps*** in the sense that there is a $j \in (0,1)$ such that, except possibly in an evanescent set,

$$\sum_\eta L_\eta \cdot |\Delta Z^\eta| \le j \ . \tag{5.6.2}$$

If $\boldsymbol{F}$ is markovian, $F_\eta[X] = f_\eta \circ X$ for some $f_\eta : \mathbb{R}^n \to \mathbb{R}^n$, then assumption J follows from the following requirement on the collaboration of $\mathbf{f}$ and $\Delta\boldsymbol{Z}$: if $x \ne y$ then for nearly all $\omega \in \Omega$

$$x + f_\eta(x)\Delta Z^\eta_s(\omega) \ne y + f_\eta(y)\Delta Z^\eta_s(\omega) \qquad 0 < s < \infty \ ;$$

this again will follow from the assumption that $x + f_\eta(x)z^\eta = y + f_\eta(y)z^\eta$ have a solution $\boldsymbol{z} \in \mathbb{R}^d$ only if $x = y \in \mathbb{R}^n$.

Theorem 5.6.1 *Suppose the coupling coefficients F^η of equation (5.6.1) satisfy the strong Lipschitz assumption L.*

(i) Then for nearly every $\omega \in \Omega$ all of the continuous maps Ξ^ω_t, $t \ge 0$, are injections of $\mathbb{R}^n$ into itself if and only if assumption J obtains.

(ii) If inequality (5.6.2) holds then for nearly every $\omega \in \Omega$ all of the Ξ^ω_t, $t \ge 0$, are also surjective and therefore are homeomorphisms from $\mathbb{R}^n$ onto itself.

Proof [92]. Multiplying $\boldsymbol{F}$ with a constant c and $\boldsymbol{Z}$ with $1/c$ does not alter premises nor conclusions and allows us to arrange things so that $2L \le 1$. We further reduce the problem by picking an instant u, and prove the claim only for all $t \le u$. Since u is arbitrary this suffices to prove both claims for all t. Since a change of probability to an equivalent one alters neither assumptions nor conclusions, Theorem 4.1.2 (ii) permits us to assume also that $\boldsymbol{Z} = \boldsymbol{Z}^u$ is an L^q-integrator with $\boldsymbol{Z}_0 = 0$ for some $q > 6n$. We also fix the exponent $p = q$.

For technical reasons we choose for the controller of $\boldsymbol{Z}$, which is the main ingredient in the definition of the norms $\| \ \|^\star_{p,M}$, not the minimal controller $\Lambda^{\langle q\rangle}[\boldsymbol{Z}]$ of theorem 4.5.1 (ii), but the controller $\Lambda^{\langle q\rangle}\big[Z^\eta, [Z^\eta, Z^\theta]/(1-j)^2\big]$. (If we are interested only in part (i), we take $j = 0$.)

First the injectivity (i). Let $x \ne y \in \mathbb{R}^n$ and set $H = H^{xy} \stackrel{\text{def}}{=} X^x - X^y$. According to equation (5.2.32), H satisfies the stochastic differential equation

$$H = (x-y) + G_\eta[H]{*}Z^\eta \ ,$$

where $$G_\eta[H] \stackrel{\text{def}}{=} F_\eta[H + X^y] - F_\eta[X^y] \ , \qquad \eta = 1,\ldots,d \ ,$$

has $G_\eta[0] = 0$ and is strongly Lipschitz with constant L_η. Let us set

$$N = N^{xy} \stackrel{\text{def}}{=} |H|^2 \quad \text{and} \quad Q = Q^{xy} \stackrel{\text{def}}{=} N^{-1}_{\boldsymbol{.}-}{*}N$$

and compute:

$$N = |x-y|^2 + 2\big\langle H_{\boldsymbol{.}-} \ast H\big\rangle + \textstyle\sum_{\nu=1}^n [H^\nu, H^\nu]$$

$$= |x-y|^2 + 2\langle H|G_\eta[H]\rangle_{.-}*Z^\eta + \langle G_\eta[H]|G_\theta[H]\rangle_{.-}*[Z^\eta, Z^\theta] \ ,$$

and $$Q = J_{\eta .-}*Z^\eta + K_{\theta\eta .-}*[Z^\theta, Z^\eta] \ , \tag{5.6.3}$$

where $J_\eta \stackrel{\text{def}}{=} \dfrac{2\langle H|G_\eta[H]\rangle}{|H|^2}$ and $K_{\theta\eta} \stackrel{\text{def}}{=} \dfrac{\langle G_\eta[H]|G_\theta[H]\rangle}{|H|^2}$

are bounded by 1 in absolute value. For this calculation to make sense $N_{.-}^{-1}$ must be N-integrable. To show that it is we *define* Q by equation (5.6.3) and note that then $dN = N_{.-}dQ$. In other words, $N = |x-y|^2 \cdot \mathscr{E}[Q]$, $\mathscr{E}[Q]$ denoting the Doléans–Dade exponential of Q. Let $S \stackrel{\text{def}}{=} \inf\{s : \Delta Q_s = -1\}$. On $[\![S]\!]$ we have $\Delta N_S = -N_{S-}$ and thus $N_S = 0$ and $X_S^x = X_S^y$. Assumption J implies $[\![S]\!] \dot\subset [X_{.-}^x = X_{.-}^y] = [N_{.-} = 0]$, which is to say that $N_{S-} = 0$ nearly on $[S < \infty]$. Now by exercise 3.9.4 (ii), $N_{S-} > 0$ on $[S < u]$ nearly, and therefore $[S < u]$ is nearly empty; $t \mapsto N_{t-}^u(\omega)$ is bounded away from zero on $[0, u]$, nearly; and by theorem 3.7.17 $N_{.-}^{-1}$ is N-integrable.

Set $$U \stackrel{\text{def}}{=} \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x \neq y\}$$

and let us remove from Ω the set

$$\bigcup \left\{ \left[\inf_{0\le t\le u} N_t^{xy} = 0\right] : (x,y) \in U \ , \ x, y \ \text{ rational } \right\} ,$$

which we now know to be nearly empty. For a fixed $\omega \in \Omega$ and $K \in \mathbb{N}$ consider the set

$$U_K(\omega) \stackrel{\text{def}}{=} \{(x,y) \in U : \inf_{0\le t\le u} N_t^{xy}(\omega) > 1/K\} \ .$$

Since $(x,y) \mapsto N_.^{xy}(\omega)$ is a continuous map from U to the càdlàg paths on $[0, u]$ given the topology of uniform convergence, $U_K(\omega)$ is open. Since the open set $\bigcup_K U_K(\omega)$ contains all rational points of U, it actually equals U. That is to say, $N_t^{xy}(\omega) > 0$ for all $\omega \in \Omega$ and all $t \le u$ simultaneously: the Ξ_t^ω are indeed injective, for all $\omega \in \Omega$ and all $t \in [0, u]$.

An aside for later use: since Λ was chosen so as to turn out to be a controller for $Q = Q^{xy}$ at this point, inequality (5.2.23) on page 290 yields $\left\| |X^x - X^y|^2 \right\|_{p,M}^\star \le |x-y|^2/(1-\gamma)$ for some suitable M and $\gamma < 1$, which reads

$$\left\| |X^x - X^y| \right\|_{p/2,M}^\star \le C^+ \cdot |x-y| \tag{+}$$

for some constant $C^+ = C_{p,n,\gamma}^+$ that is independent of x, y.

We turn to the surjectivity (ii) of the Ξ_t^ω, assuming inequality (5.6.2). Set

$$\overline{Q}_t = \overline{Q}_t^{xy} \stackrel{\text{def}}{=} \int_0^t \frac{d[Q,Q]}{1+\Delta Q} - Q_t = \sum_{0<s\le t} \frac{(\Delta Q_s)^2}{1+\Delta Q_s} - Q_t$$

The integral has to understood in the pathwise Stieltjes sense, of course. An easy calculation gives $Q + \overline{Q} + [Q, \overline{Q}] = 0$. By exercise 3.9.3, $\mathscr{E}[Q]$ and $\mathscr{E}[\overline{Q}]$ are inverses of each other, and so

$$H^{-1} = |X^x - X^y|^{-2} = |x-y|^{-2}\mathscr{E}[\overline{Q}] \ .$$

We will show next that Λ is a controller for $\overline{Q}$. To this end we claim that

$$\Delta Q_s \geq -1 + (1-q)^2 \quad \text{and thus} \quad \frac{1}{1+\Delta Q} \leq \frac{1}{(1-q)^2} \ .$$

Indeed, on $[\Delta Q_s < -1 + (1-j)^2]$ we have

$$\Delta Q = 2\frac{\langle H_{\bullet-}|G_\eta[H]_{\bullet-}\Delta Z^\eta\rangle}{|H|^2_{\bullet-}} + \frac{\langle G_\eta[H]_{\bullet-}\Delta Z^\eta|G_\theta[H]_{\bullet-}\Delta Z^\theta\rangle}{|H|^2_{\bullet-}} < -1 + (1-j)^2$$

$$\Longrightarrow \quad 2\langle H_{\bullet-}|G_\eta[H]_{\bullet-}\Delta Z^\eta\rangle + \langle G_\eta[H]_{\bullet-}\Delta Z^\eta|G_\theta[H]_{\bullet-}\Delta Z^\theta\rangle + |H|^2_{\bullet-}$$
$$< (1-j)^2|H|^2_{\bullet-}$$
$$\Longrightarrow \quad \big|H_{\bullet-} + G_\eta[H]_{\bullet-}\Delta Z^\eta\big|^2 < (1-j)^2|H|^2_{\bullet-}$$
$$\Longrightarrow \quad \big|H_{\bullet-} + G_\eta[H]_{\bullet-}\Delta Z^\eta\big| < (1-j)|H|_{\bullet-}$$
$$\Longrightarrow \quad \big|G_\eta[H]_{\bullet-}\Delta Z^\eta\big| > j|H|_{\bullet-} \qquad \Longrightarrow \qquad \big|L_\eta \Delta Z^\eta\big| > j \ .$$

Due to the assumption (5.6.2), $[\Delta Q_s < -1 + (1-j)^2]$ is evanescent.

Our choice of the controller Λ was made precisely to assure that it is also a controller for $\overline{Q}^{xy}$. The argument leading to (+) applies and gives here

$$\big\| |X^x - X^y|^{-1} \big\|^\star_{p/2,M} \leq C^- \cdot |x-y|^{-1} \ , \tag{$-$}$$

with constant C^- independent of x, y. Set now

$$Y^x \stackrel{\text{def}}{=} \begin{cases} \big|X^{x/|x|^2} - X^0\big|^{-1} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0. \end{cases}$$

Then $$\big|Y^x - Y^y\big| \leq \frac{\big|X^{x/|x|^2} - X^{y/|y|^2}\big|}{\big|X^{x/|x|^2} - X^0\big|\big|X^{y/|y|^2} - X^0\big|}$$

and $$\big\| \big|Y^x - Y^y\big| \big\|^\star_{p/6,M} \leq C^-(C^+)^2 \cdot |x||y| \left| \frac{x}{|x|^2} - \frac{y}{|y|^2} \right| = C \cdot |x-y| \ .$$

We use corollary 5.2.23 on page 295 to show that after tossing out another nearly empty set, a version of $Y^x(\omega)$ can be chosen that is continuous in x for all $\omega \in \Omega$, in particular at $x = 0$. This means that

$$\lim_{|x|\to\infty} \big|X^x_t(\omega)\big| = \infty \ :$$

$x \mapsto X^x_t(\omega)$ maps the point at infinity in $\mathbb{R}^n$ to itself, for all t and all $\omega \in \Omega$. Thus for such ω and t, $\Xi^\omega_t : x \mapsto X^x_t(\omega)$ can be viewed as a continuous injection of the n-sphere into itself. By Brouwer's invariance–of–domain theorem, the injectivity implies that its image is open; the compactness of the sphere implies that this image is closed; the connectivity of the n-sphere implies that the map in question is surjective: it is a homeomorphism. ∎

Exercise 5.6.2 Let $Y = Y^\mu_\nu$ be an $n \times n$-matrix of L^q-integrators, $q \geq 2$. Consider $Y_t(\omega)$ as a linear operator from euclidean space $\mathbb{R}^n$ to itself, with operator norm $\|Y\|$. Its jump is the matrix

$$\Delta Y_s \overset{\text{def}}{=} (\Delta Y^\mu_{\nu\, s})^{\mu=1\ldots n}_{\nu=1\ldots n} \ ;$$

its square function is[1]

$$[Y,Y] = \big([Y,Y]\big)^\mu_\nu \overset{\text{def}}{=} [Y^\mu_\rho, Y^\rho_\nu] \ ,$$

which by theorems 3.8.4 and 3.8.9 is an $L^{q/2}$-integrator. Set

$$\overline{Y}_t \overset{\text{def}}{=} -Y_t + {}^c[Y,Y]_t + \sum_{0<s\leq t} (I + \Delta Y_s)^{-1} (\Delta Y_s)^2 \ .$$

(i) Assume that $\overline{Y}$ is an L^0-integrator. Then the solutions of

$$D_t = I + \int_0^t D_{s-}\, dY_s \quad \text{and} \quad \overline{D}_t = I + \int_0^t d\overline{Y}\, \overline{D}_{s-}$$

are inverse to each other:

$$D_t \overline{D}_t = I \qquad \forall\, t \geq 0 \ .$$

Here[1]

$$(d\overline{Y}\,\overline{D}_{\bullet -})^\mu_\nu \overset{\text{def}}{=} D^\rho_{\nu\, \bullet -}\, d\overline{Y}^\mu_\rho \ .$$

(ii) If $\sup_{(s,\omega)\in \boldsymbol{B}} \|\Delta Y_s(\omega)\| < 1$, then $(I + \Delta Y_s)^{-1}$ is bounded. If $(I + \Delta Y_s)^{-1}$ is bounded, then

$$J_t \overset{\text{def}}{=} \sum_{0<s\leq t} (I + \Delta Y_s)^{-1} (\Delta Y_s)^2$$

is an $L^{q/2}$-integrator, and so is $\overline{Y}$.

Markovian Stochastic Flows

If the coupling coefficient of (5.6.1) is markovian, that equation becomes

$$X^x_t = x + \int_0^t \boldsymbol{f}(X^x_{s-})\, d\boldsymbol{Z}_s \ . \tag{5.6.4}$$

The f_η comprising $\boldsymbol{f}$ are assumed Lipschitz with constant L, and the driver $\boldsymbol{Z}$ is an arbitrary L^0-integrator with $\boldsymbol{Z}_0 = 0$, at least for a while. Summary 5.4.7 on page 317 provides the universal solution $\overline{X}(x, \boldsymbol{z}_\bullet)$ of

$$\overline{X}_t = x + \int_0^t \boldsymbol{f}(\overline{X}_{s-})\, d\overline{\boldsymbol{Z}}_s \ . \tag{5.6.5}$$

Since at this point we are interested only in constant initial conditions, we restrict $\overline{X}$ to $\mathbb{R}^n \times \mathscr{D}^d$, so that it becomes a map from $\mathbb{R}^n \times \mathscr{D}^d$ to $\mathscr{D}^n$, adapted to the filtrations $\mathcal{B}^\bullet(\mathbb{R}^n)\otimes\mathcal{F}_{\bullet}[\mathscr{D}^d]$ and $\mathcal{F}_{\bullet}[\mathscr{D}^n]$ (see item 2.3.8). $\overline{\boldsymbol{Z}}$ is the process representing $\boldsymbol{Z}$ on the path space $\mathbb{R}^n\times\mathscr{D}^d$: $\overline{\boldsymbol{Z}}_t(x,\boldsymbol{z}_{\bullet})=\boldsymbol{z}_t$, and the assumption $\boldsymbol{Z}_0=0$ has the effect that any of the probabilities of $\overline{\mathfrak{P}}\overset{\text{def}}{=}\underline{\boldsymbol{Z}}[\mathfrak{P}]\subset\mathfrak{P}[\overline{\boldsymbol{Z}}]$ is carried by the set $[\overline{\boldsymbol{Z}}_0=0]=\{(x,\boldsymbol{z}_{\bullet}):\boldsymbol{z}_0=0\}$. Taking for the parameter domain U of theorem 5.2.24 the space $\mathbb{R}^n$ itself, we see that $\overline{X}$ can be modified on a $\mathfrak{P}[\overline{\boldsymbol{Z}}]$-nearly empty set in such a way that both

$$\boldsymbol{z}_{\bullet}\mapsto\overline{X}_{\bullet}(x,\boldsymbol{z}_{\bullet})\ \text{ is adapted to } \mathcal{F}_{\bullet}[\mathscr{D}^d] \text{ and } \mathcal{F}_{\bullet}[\mathscr{D}^n] \quad \forall\, x\in\mathbb{R}^n \tag{5.6.6}$$

and $$x\mapsto\overline{X}_{\bullet}(x,\boldsymbol{z}_{\bullet})\ \text{ is continuous}^{21}\ \text{ from } \mathbb{R}^n \text{ to } \mathscr{D}^n \quad \forall\, \boldsymbol{z}_{\bullet}\in\mathscr{D}^d. \tag{5.6.7}$$

Note that $\overline{X}$ is a construct made from the coupling coefficient $\boldsymbol{f}$ alone; in particular, the prevailing probability does not enter its definition. $X^x_{\bullet}\overset{\text{def}}{=}\overline{X}_{\bullet}(x,\boldsymbol{Z}_{\bullet})$ is a solution of (5.6.4), and any other solution that depends continuously on x differs from $X^x_{\bullet}$ in a nearly empty set that does not depend on x.

Here is a straightforward observation. Let S be a stopping time and $t\ge 0$. Then

$$\begin{aligned} X^x_{S+t} &= X^x_S+\int_S^{S+t}\boldsymbol{f}\big(X^x_{\sigma-}\big)\,d(\boldsymbol{Z}_\sigma-\boldsymbol{Z}_S)\\ &= X^x_S+\int_0^t\boldsymbol{f}\big(X^x_{(S+\tau)-}\big)\,d(\boldsymbol{Z}_{S+\tau}-\boldsymbol{Z}_S)\,. \end{aligned}$$

This says that the process $X_{S+\bullet}$ satisfies the same differential equation (5.6.4) as does $X_{\bullet}$, except that the initial condition is X_S and the driver $\boldsymbol{Z}_{\bullet}-\boldsymbol{Z}_S$. Now upon representing this driver on $\mathscr{D}^d$, every $\mathbb{P}\in\mathfrak{P}[\boldsymbol{Z}]$ turns into a probability in $\mathfrak{P}[\overline{\boldsymbol{Z}}]$, inasmuch as $\boldsymbol{Z}_{\bullet}-\boldsymbol{Z}_S$ is a $\mathbb{P}$-integrator. Therefore this stochastic differential equation is solved by

$$X^x_{S+t}=\overline{X}_t(X^x_S,\boldsymbol{Z}_{S+\bullet}-\boldsymbol{Z}_S)\,. \tag{5.6.8}$$

Applying this very argument to equation (5.6.5) on $\mathbb{R}^n\times\mathscr{D}^d$ results in

$$\overline{X}_{S+t}(x,\boldsymbol{z}_{\bullet})=\overline{X}_t\big(\overline{X}_S(x,\boldsymbol{z}_{\bullet}),\boldsymbol{z}_{S+\bullet}-\boldsymbol{z}_S\big)\,. \tag{5.6.9}$$

For any $\mathcal{F}_{\bullet}[\mathscr{D}^d]$-stopping time S, any $t\ge 0$, and any $x\in\mathbb{R}^n$, this equation holds a priori only $\mathfrak{P}[\overline{\boldsymbol{Z}}]$-nearly, of course. At a fixed stopping time S we may assume, by throwing out a $\mathfrak{P}[\overline{\boldsymbol{Z}}]$-nearly empty set, that (5.6.9) holds for all rational instants t and all rational points $x\in\mathbb{R}^n$. Since $\overline{X}$ is continuous in its first argument, though, and $t\mapsto\overline{X}_t(x,\boldsymbol{z}_{\bullet})$ is right-continuous, we get

Proposition 5.6.3 *The universal solution for equation (5.6.4) satisfies (5.6.6) and (5.6.7); and for any stopping time S there exists a nearly empty set outside which equation (5.6.9) holds simultaneously for all $x\in\mathbb{R}^n$ and all $t\ge 0$.*

Problem 5.6.4 Can (5.6.9) be made to hold identically for all $s\ge 0$?

Markovian Stochastic Flows Driven by a Lévy Process

The situation and notations are the same as in the previous subsection, except that we now investigate the case that the driver $\boldsymbol{Z}$ is a Lévy process. In this case THE previsible controller $\Lambda_t[\boldsymbol{Z}]$ is just a constant multiple of time t (equation (4.6.30)) and THE time transformation $T^{\bullet}$ is also simply a linear scaling of time. Let us define the positive linear operator T_t on $C_b(\mathbb{R}^n)$ by

$$T_t\phi(x) \stackrel{\text{def}}{=} \mathbb{E}[\phi(X_t^x)] = \mathbb{E}[\phi \circ \overline{X}_t(x, \boldsymbol{Z}_{\bullet})] \ . \tag{5.6.10}$$

It follows from inequality (5.2.36) that $T_t\phi$ is continuous; and it is not too hard to see with the help of inequality (5.2.25) that T_t maps $C_0(\mathbb{R}^n)$ into $C_0(\mathbb{R}^n)$ when $\boldsymbol{f}$ is bounded. Now fix an $\mathcal{F}_{\bullet}$-stopping time S. Since $\boldsymbol{Z}_{\bullet+S}-\boldsymbol{Z}_S$ is independent of $\mathcal{F}_S$ and has the same law as $\boldsymbol{Z}_{\bullet}$ (exercise 4.6.1),

$$\begin{aligned}\mathbb{E}[\phi \circ \overline{X}_t(x, \boldsymbol{Z}_{\bullet+S} - \boldsymbol{Z}_S)|\mathcal{F}_S] &= \mathbb{E}[\phi \circ \overline{X}_t(x, \boldsymbol{Z}_{\bullet+S} - \boldsymbol{Z}_S)] \\ &= \mathbb{E}[\phi \circ \overline{X}_t(x, \boldsymbol{Z}_{\bullet})] = T_t\phi(x) \ ,\end{aligned}$$

whence

$$\mathbb{E}\big[\phi \circ \overline{X}_t(X_S^x, \boldsymbol{Z}_{\bullet+S} - \boldsymbol{Z}_S)|\mathcal{F}_S\big] = T_t\phi \circ X_S^x \ ,$$

thanks to exercise A.3.26 on page 408. With equation (5.6.8) this gives

$$\mathbb{E}[\phi \circ X_{S+t}^x|\mathcal{F}_S] = T_t\phi \circ X_S^x \tag{5.6.11}$$

and, taking the expectation,

$$T_{s+t}\phi(x) = T_s[T_t\phi](x) \tag{5.6.12}$$

for $s, t \in \mathbb{R}_+$ and $x \in \mathbb{R}^n$. The remainder of this subsection is given over to a discussion of these two equalities.

5.6.5 (The Markov Property) Taking in (5.6.11) the conditional expectation under X_S^x (see page 407) gives

$$\mathbb{E}\big[\phi \circ X_{S+t}^x|\mathcal{F}_S\big] = \mathbb{E}\big[\phi \circ X_{S+t}^x|X_S^x\big] \circ X_S^x \ . \tag{5.6.13}$$

That is to say, as far as the distribution of X_{S+t}^x is concerned, knowledge of the whole path up to time S provides no better clue than knowing the position X_S^x at that time: "the process continually forgets its past." A process satisfying (5.6.13) is said to have the ***strong Markov property***. If (5.6.13) can be ascertained only at deterministic instants S, then the process has the "plain" ***Markov property***. Equation (5.6.11) is actually stronger than the strong Markov property, in this sense: the function $T_t\phi$ is a common conditional expectation of $\phi(X_{S+t}^x)$ under X_S^x, one that depends neither on S nor on x. We leave it to the reader to parlay equation (5.6.11) into the following. Let $F : \mathscr{D}^n \to \mathbb{R}$ be bounded and measurable on the Baire σ-algebra for the pointwise topology, that is, on $\mathcal{F}_\infty^0[\mathscr{D}^n]$. Then

$$\mathbb{E}\big[F(X_{S+\bullet}^x)|\mathcal{F}_S\big] = \mathbb{E}\big[F(X_{S+\bullet}^x)|X_S^x\big] \circ X_S^x \ . \tag{5.6.14}$$

To paraphrase: "in order to predict at time S anything[45] about the future of $X^x_.$, knowledge of the whole history $\mathcal{F}_S$ of the world up to that time is not more helpful than knowing merely the position X^x_S at that time."

Exercise 5.6.6 Let us wean the processes X^x, $x \in \mathbb{R}^n$, of their provenance, by representing each on $\mathscr{D}^n$ (see items 2.3.8–2.3.11). The image under $\underline{X}^x$ of the given probability is a probability on $\mathcal{F}_\infty[\mathscr{D}^n]$, which shall be written $\mathbb{P}^x$; the corresponding expectation is $\mathbb{E}^x$. The evaluation process $(t, x_.) \mapsto x_t$ will be written $\overline{X}$, as often before. Show the following. (i) Under $\mathbb{P}^x$, $\overline{X}$ starts at x: $\mathbb{P}^x[\overline{X}_0 = x] = 1$. (ii) For $F \in L^\infty(\mathcal{F}_\infty[\mathscr{D}^n])$, $x \mapsto \mathbb{E}^x[F]$ is universally measurable. (iii) For all $\phi \in C_0(E)$, all stopping times S on $\mathcal{F}_.[\mathscr{D}^n]$, all $t \geq 0$, and all $x \in E$ we have $\mathbb{P}^x$-almost surely $\mathbb{E}^x[\phi(X_{S+t})|\mathcal{F}_S] = T_t\phi \circ X_S$. (iv) $\overline{X}$ is quasi-left-continuous.

5.6.7 (The Feller Semigroup of the Flow) Equation (5.6.12) says that the operators T_t form a semigroup under composition: $T_{s+t} = T_s \circ T_t$ for $s, t \geq 0$. Since evidently $\sup\{T_t\phi\,(x) : \phi \in C_0(E)\,,\ 0 \leq \phi \leq 1\} = \mathbb{E}[1] = 1$, we have

Proposition 5.6.8 *$\{T_t\}_{t\geq 0}$ forms a conservative Feller semigroup on $C_0(\mathbb{R}^n)$.*

Let us go after the generator of this semigroup. Itô's formula applied to $X^x_.$ and a function ϕ on $\mathbb{R}^n$ of class[10] C^2_b gives[1]

$$
\begin{aligned}
\phi(X^x_t) &= \phi(X^x_0) + \int_0^t \phi_{;\nu}(X^x_{s-})\, dX^x_s + \frac{1}{2}\int_0^t \phi_{;\mu\nu}(X^x_s) d^c[X^{x\mu}, X^{x\nu}] \\
&\quad + \textstyle\sum_{0\leq s\leq t} \phi\big(X^x_{s-} + \Delta X^x_s\big) - \phi(X^x_{s-}) - \phi_{;\nu}(X^x_{s-})\Delta X^{x\nu}_s \\
&= \phi(x) + \int_0^t \big(\phi_{;\nu} f^\nu_\eta\big)(X^x_{s-})\, dZ^\eta_s + \frac{1}{2}\int_0^t \big(\phi_{;\mu\nu} f^\mu_\eta f^\nu_\theta\big)(X^x_s)\, d^c[Z^\eta, Z^\theta]_s \\
&\quad + \textstyle\sum_{0\leq s\leq t} \phi\big(X^x_{s-} + f_\eta(X^x_{s-})\Delta Z^\eta_s\big) - \phi(X^x_{s-}) - \big(\phi_{;\nu} f^\nu_\eta\big)(X^x_{s-})\Delta Z^\eta_s \\
&= \phi(x) + \int_0^t \big(\phi_{;\nu} f^\nu_\eta\big)(X^x_{s-})\, \big(dZ^\eta_s - y^\eta[|\boldsymbol{y}| > 1]\, \jmath_{\boldsymbol{Z}}(d\boldsymbol{y}, ds)\big) \\
&\quad + \frac{1}{2}\int_0^t \big(\phi_{;\mu\nu} f^\mu_\eta f^\nu_\theta\big)(X^x_s)\, d^c[Z^\eta, Z^\theta]_s \\
&\quad + \int_0^t \phi\big(X^x_{s-} + f_\eta(X^x_{s-})y^\eta\big) - \phi(X^x_{s-}) - \big(\phi_{;\nu} f^\nu_\eta\big)(X^x_{s-})y^\eta[|\boldsymbol{y}|\leq 1]\, \jmath_{\boldsymbol{Z}}(d\boldsymbol{y}, ds) \\
&= Mart_t + \phi(x) + \int_0^t \big(\phi_{;\nu} f^\nu_\eta\big)(X^x_s) A^\eta\, ds + \frac{1}{2}\int_0^t \big(\phi_{;\mu\nu} f^\mu_\eta f^\nu_\theta\big)(X^x_s) B^{\eta\theta}\, ds \\
&\quad + \int_0^t \int \phi\big(X^x_s + f_\eta(X^x_s)y^\eta\big) - \phi(X^x_s) - \big(\phi_{;\mu} f^\mu_\eta\big)(X^x_s)y^\eta[|\boldsymbol{y}|\leq 1]\, \nu(d\boldsymbol{y})\, ds
\end{aligned}
$$

In the penultimate equality the large jumps were shifted into the first term as in equation (4.6.32). We take the expectation and differentiate in t at $t = 0$ to obtain

[45] Well, at least anything that depends Baire measurably on the path $t \mapsto X^x_{S+t}$.

Proposition 5.6.9 *The generator $\mathcal{A}$ of $\{T_t\}_{t\ge 0}$ acts on a function $\phi \in C_0^2$ by*[1]

$$\mathcal{A}\phi(x) = A^\eta f_\eta^\nu(x)\frac{\partial\phi}{\partial x^\eta}(x) + \frac{B^{\eta\theta}}{2} f_\eta^\mu(x) f_\theta^\nu(x)\frac{\partial^2\phi}{\partial x^\eta\partial x^\theta}(x)$$
$$+ \int \Big(\phi\big(x + f_\eta(x)y^\eta\big) - \phi(x) - \frac{\partial\phi}{\partial x^\mu}(x)\, f_\eta^\mu(x)\, y^\eta[|\boldsymbol{y}| \le 1]\Big)\nu(dy)\ .$$

Differentiating at $t \ne 0$ gives

$$\frac{dT_t\phi}{dt} = T_t\mathcal{A}\phi = \overline{\mathcal{A}}T_t\phi\ ,$$

where $\overline{\mathcal{A}}$ is the closure of $\mathcal{A}$. Suppose the coefficients $\boldsymbol{f}$ have two continuous bounded derivatives. Then, by theorem 5.3.16, $x \mapsto X_t^x$ is twice continuously differentiable as a map from $\mathbb{R}^n$ to any of the L^p, $p < \infty$, and hence $x \mapsto T_t\phi(x) = \mathbb{E}[\phi(X_t^x)]$ is twice continuously differentiable. On this function $\overline{\mathcal{A}}$ agrees with $\mathcal{A}$:

Corollary 5.6.10 *If $\boldsymbol{f} \in C_b^2$ and $\phi \in C_0^2$, then $u(t,x) \stackrel{\text{def}}{=} \mathbb{E}[\phi(X_t^x)]$ is continuous and twice continuously differentiable in x and satisfies the evolution equation*

$$\frac{du(t,x)}{dt} = \mathcal{A}u(t,x)\ .$$

5.7 Semigroups, Markov Processes, and PDE

We have encountered several occasion where a Feller semigroup arose from a process (page 268) or a stochastic differential equation (item 5.6.7) and where a PDE appeared in such contexts (item 5.5.8, exercise 5.5.13, corollary 5.6.10). This section contains some rudimentary discussions of these connections.

Stochastic Representation of Feller Semigroups

Not only do some processes give rise to Feller semigroups, every Feller semigroup comes this way:

Definition 5.7.1 *A **stochastic representation** of the conservative Feller semigroup $T_.$ consists of a filtered measurable space $\big(\Omega, \mathcal{F}_.\big)$ together with an E-valued adapted*[46] *process $X_.$ and a slew $\{\mathbb{P}^x\}$ of probabilities on $\mathcal{F}_\infty$, one for every point $x \in E$, satisfying the following description:*
(i) for every $x \in E$, X starts at x under $\mathbb{P}^x$: $\mathbb{P}^x[X_0 = x] = 1$;
(ii) $x \mapsto \mathbb{E}^x[F] \stackrel{\text{def}}{=} \int F\, d\mathbb{P}^x$ is universally measurable, for every $F \in L^\infty(\mathcal{F}_\infty)$;
(iii) for all $\phi \in C_0(E)$, $0 \le s < t < \infty$, and $x \in E$ we have $\mathbb{P}^x$-almost surely

$$\mathbb{E}^x[\phi(X_t)|\mathcal{F}_s] = T_{t-s}\phi \circ X_s\ . \tag{5.7.1}$$

[46] X_t is $\mathcal{F}_t/\mathcal{B}^*(E)$-measurable for $0 \le t < \infty$ — see page 391.

5.7.2 Here are some easily verified consequences. (i) With $s = 0$ equation (5.7.1) yields

$$\mathbb{E}^x\big[\phi(X_t)\big] = T_t\phi\,(x)\ .$$

(ii) For any $\phi \in C_0(E)$ and $x \in E$, $t \mapsto \phi(X_t)$ is a uniformly continuous curve in the complete seminormed space $\mathcal{L}^2(\mathbb{P}^x)$. So is the curve $t \mapsto T_{u-t}\phi\,(X_t)$ for $0 \le t \le u$. (iii) For any $\alpha > 0$, $x \in E$, and positive $\gamma \in C_0(E)$,

$$t \mapsto Z_t^{\alpha,\gamma} \stackrel{\text{def}}{=} e^{-\alpha t} \cdot U_\alpha\gamma \circ X_t$$

is a positive bounded $\mathbb{P}^x$-supermartingale on $\big(\Omega, \mathcal{F}_{\bullet}\big)$. Here $U_{\bullet}$ is the resolvent, see page 463.

Theorem 5.7.3 *(i) Every conservative Feller semigroup $T_{\bullet}$ has a stochastic representation. (ii) In fact, there is a stochastic representation $\big(\Omega, \mathcal{F}_{\bullet}, X_{\bullet}, \mathfrak{P}\big)$ in which $\mathcal{F}_{\bullet}$ satisfies the natural conditions*[47] *with respect to $\mathfrak{P} \stackrel{\text{def}}{=} \{\mathbb{P}^x\}_{x\in E}$ and in which the paths of $X_{\bullet}$ are right-continuous with left limits and stay in compact subsets of E during any finite interval. Such we call a* ***regular stochastic representation****.*

(iii) A regular stochastic representation has these additional properties: ***The Strong Markov Property****: for any finite $\mathcal{F}_{\bullet}$-stopping time S, number $\sigma \ge 0$, bounded Baire function f, and $x \in E$ we have $\mathbb{P}^x$-almost surely*

$$\mathbb{E}^x\big[f(X_{S+\sigma})|\mathcal{F}_S\big] = \int_E T_\sigma(X_S, dy)\ f(y) = \mathbb{E}^{X_S}\big[f(X_\sigma)\big]\ . \tag{5.7.2}$$

Quasi-Left-Continuity*: for any strictly increasing sequence T_n of $\mathcal{F}_{\bullet}$-stopping times with finite supremum T and all $\mathbb{P}^x \in \mathfrak{P}$*

$$\lim X_{T_n} = X_T \qquad\qquad \mathbb{P}^x\text{-nearly.}$$

Blumenthal's Zero-One Law*: the $\mathfrak{P}$-regularization of the basic filtration of $X_{\bullet}$ is right-continuous.*

Remarks 5.7.4 Equation (5.7.2) has this consequence:[48] under $\mathbb{P} = \mathbb{P}^x$

$$\mathbb{E}^{\mathbb{P}}[f \circ X_{S+\sigma}|\mathcal{F}_S] = \mathbb{E}^{\mathbb{P}}[f \circ X_{S+\sigma}|X_S] \circ X_S\ . \tag{5.7.3}$$

This says that as far as the distribution of $X_{S+\sigma}$ is concerned, knowledge of the whole path up to time S provides no better clue than knowing the position X_S at that time: "the process continually forgets its past." An adapted process X on $(\Omega, \mathcal{F}_{\bullet}, \mathbb{P})$ obeying (5.7.3) for all finite stopping times S and $\sigma \ge 0$ is said to have the ***strong Markov property***. It has the "plain" ***Markov property*** as long as (5.7.3) can be ascertained for sure times S.

---

[47] See definition 1.3.38 and warning 1.3.39 on page 39.

[48] See theorem A.3.24 for the conditional expectation under the map X_S. $\mathbb{E}[f \circ X_T|X_S]$ is a function on the range of X_S and is unique up to sets negligible for the law of X_S.

Actually, equation (5.7.2) gives much more than merely the strong Markov property of $X_.$ on $(\Omega, \mathcal{F}_., \mathbb{P}^x)$ for every starting point x. Namely, it says that the Baire function $x' \mapsto \int T_\sigma(x', dy)\, f(y)$ serves as a common member of every one of the classes[48] $\mathbb{E}^x\big[f(X_{S+\sigma})|X_S\big]$. Note that this function depends neither on S nor on x.

Consider the case that S is an instant s and apply (5.7.2) to a function $\phi \in C_0(E)$ and a Borel set B in E.[49] Then

$$\mathbb{E}^x\big[f \circ X_{s+\sigma}|\mathcal{F}_s\big] = T_\sigma\phi \circ X_s$$

and
$$\mathbb{P}^x\big[X_{s+\sigma} \in B\big|\mathcal{F}_s\big] = T_\sigma(X_s, B) \ . \tag{5.7.4}$$

Visualize X as the position of a meandering particle. Then (5.7.4) says that the probability of finding it in B at time $s+\sigma$ given the whole history up to time s is the same as the probability that during σ it made its way from its position at time s to B, *no matter how it got to that position.* ∎

Exercise 5.7.5 [11, page 50] If for every compact $K \subset E$ and neighborhood G of K

$$\lim_{t\to 0} \sup_{x\in K} \frac{T_t(x, E\setminus G)}{t} = 0 \ ,$$

then the paths of $X_.$ are $\mathbb{P}^x$-almost surely continuous, for all $x \in E$.

Exercise 5.7.6 For every $x \in E$ and every function ψ in the domain $\breve{\mathcal{D}}$ of the natural extension $\breve{A}$ of the generator (see pages 467–468)

$$M_t^\psi \stackrel{\text{def}}{=} \psi(X_t) - \int_0^t \mathcal{A}\psi \circ X_s\, ds$$

is a $\mathbb{P}^x$-martingale. Conversely, if there exists a $\phi \in \breve{C}$ so that for all $x \in E$ $M_t^\psi \stackrel{\text{def}}{=} \psi(X_t) - \int_0^t \phi(X_s)\, ds$ is a $\mathbb{P}^x$-martingale, then $\psi \in \breve{\mathcal{D}}$ and $\breve{A}\psi = \phi$.

Proof of Theorem 5.7.3. To start with, let us assume that E is compact.

(i) Prepare, for every $t \in \mathbb{R}_+$, a copy E_t of E and take for Ω the product

$$E^{[0,\infty)} = \prod_{t\in[0,\infty)} E_t$$

of all $[0,\infty)$-tuples $(x_t)_{0\le t<\infty}$ with entry x_t in E_t. Ω is compact in the product topology. X_t is simply the t^{th} coordinate: $X_t\big((x_s)_{0\le s<\infty}\big) = x_t$. Next let $\mathbb{T}$ denote the collection of all finite ordered subsets $\tau \subset [0,\infty)$ and $\mathbb{T}_0$ those that contain $t_0 = 0$; both $\mathbb{T}$ and $\mathbb{T}_0$ are ordered by inclusion. For any $\tau = \{t_1 \ldots t_k\} \in \mathbb{T}$ (with $t_i < t_{i+1}$) write

$$E_\tau = E_{t_1\ldots t_k} \stackrel{\text{def}}{=} E_{t_1} \times \cdots \times E_{t_k} \ .$$

There are the natural projections

$$\begin{aligned} X_\tau = X_{t_1\ldots t_k} : \quad & E^{[0,\infty)} \quad \to \quad E_\tau \\ & : (x_t)_{0\le t<\infty} \mapsto (x_{t_1}, \ldots, x_{t_k}) \ . \end{aligned}$$

---

[49] See convention A.1.5 on page 364.

"X_τ forgets the coordinates not listed in τ." If τ is a singleton $\{t\}$, then X_τ is the t^{th} coordinate function X_t, so that we may also write

$$X_{t_1\ldots t_k} = \big(X_{t_1}, \ldots, X_{t_k}\big) .$$

We shall define the expectations $\mathbb{E}^x$ first on a special class of functions, the ***cylinder functions***. $F : E^{[0,\infty)} \to \mathbb{R}$ is a cylinder function based on $\tau = \{t_0 t_1 \ldots t_k\} \in \mathbb{T}_0$ if there exists $f : E_\tau \to \mathbb{R}$ with $F = f \circ X_\tau$, i.e.,

$$F = f(X_{t_0 t_1 \ldots t_k}) = f(X_{t_0}, \ldots, X_{t_k}) , \qquad t_0 = 0 .$$

We say that F is Borel or continuous, etc., if f is. If F is based on τ, it is also based on any τ' containing τ; in the representation $F = f \circ X_{\tau'}$, f simply does not depend on the coordinates that are not listed in τ. For instance, if F is based on X_τ and G on $X_{\tau'}$, then both are based on X_υ, where $\upsilon \stackrel{\text{def}}{=} \tau \cup \tau'$, and can be written $F = f \circ X_\upsilon$ and $G = g \circ X_\upsilon$; thus $F + G = (f + g) \circ X_\upsilon$ is again a cylinder function. Repeating this with $\cdot, \wedge, \vee$ replacing $+$ we see that the Borel or continuous cylinder functions form both an algebra and a vector lattice.

We are ready to define $\mathbb{P}^x$. Let $F = f \circ X_{t_0 t_1 \ldots t_k}$ be a bounded Borel cylinder function. Define inductively $f^{(k)} = f$, $\sigma_i = t_i - t_{i-1}$, and

$$f^{(i-1)}(x_0, x_1, \ldots, x_{k-1}) \stackrel{\text{def}}{=} \int T_{\sigma_i}(x_{i-1}, dx_i)\, f^{(i)}(x_0, x_1, \ldots, x_{i-1}, x_i) ,$$

as long as $i \geq 1$, and finally set $\mathbb{E}^x[F] \stackrel{\text{def}}{=} f^{(0)}(x)$. In other words,

$$\mathbb{E}^x[F] \stackrel{\text{def}}{=} \int T_{t_1}(x, dx_1) \int \cdots \int T_{t_{k-1}-t_{k-2}}(x_{k-2}, dx_{k-1}) \int T_{t_k - t_{k-1}}(x_{k-1}, dx_k)$$
$$f(x, x_1, \ldots, x_{k-2}, x_{k-1}, x_k) . \tag{5.7.5}$$

To see that this makes sense assume for the moment that $f \in C(E_\tau)$. We see by inspection[50] that then $f^{(i)}$ belongs to $C(E_{t_0 \ldots t_i})$, $i = k, k-1, \ldots, 0$. Thus

$$x \mapsto \mathbb{E}^x[F] \text{ is continuous for } F \in C(E_\tau) \circ X_\tau .$$

Next consider the class of bounded Borel functions f on $E_{t_0 \ldots t_k}$ such that $f^{(i)}$ is a bounded Borel function on $E_{t_0 \ldots t_i}$ for all i. It contains $C(E_\tau)$ and is closed under bounded sequential limits, so it contains all bounded Borel functions: the iterated integral (5.7.5) makes sense.

Suppose that f does not depend on the coordinate x_j, for some j between 1 and k. The Chapman–Kolmogorov equation (A.9.9) applied with $s = \sigma_j$, $t = \sigma_{j+1}$, and $y = x_{j+1}$ to $\phi(y) = f^{(j+1)}(x_0, \ldots, x_{j-1}, y)$ shows that for

[50] This is particularly obvious when f is of the form $f(x_0, \ldots, x_k) = \prod_i \phi_i(x_i)$, $\phi_i \in C$, or is a linear combination of such functions. The latter form an algebra uniformly dense in $C(E_\tau)$ (theorem A.2.2).

$i < j$ we get the same functions $f^{(i)}$ whether we consider F as a cylinder function based on X_τ or on $X_{\tau'}$, where τ' is τ with t_j removed. Using this observation it is easy to see that $\mathbb{E}^x$ is well-defined on Borel cylinder functions and is linear. It is also evidently positive and has $\mathbb{E}^x[1] = 1$. It is a positive linear functional of mass one defined on the continuous cylinder functions, which are uniformly dense[50] in C: $\mathbb{E}^x$ is a Radon measure. Also, $\mathbb{E}^x[\phi(X_0)] = \phi(x)$. Let $\mathcal{F}^0_{\bullet}$ denote the basic filtration of $X_\bullet$. Its final σ-algebra $\mathcal{F}^0_\infty$ is clearly nothing but the Baire σ-algebra of Ω. The collection of all bounded Baire functions F on Ω for which $x \mapsto \mathbb{E}^x[F]$ is Baire measurable on E contains the cylinder functions and is sequentially closed. Thus it contains all Baire functions on Ω.

Only equation (5.7.1) is left to be proved. Fix $0 \le s < t$, and let F be a Borel cylinder function based on a partition $\tau \subset [0, s]$ in $\mathbb{T}_0$ and let $\phi \in C_0(E)$. The very definition of $\mathbb{E}^x$ gives

$$\mathbb{E}^x\big[F \cdot \phi(X_t)\big] = \mathbb{E}^x\big[F \cdot T_{t-s}\phi(X_s)\big] \; .$$

Since the functions F generate $\mathcal{F}^0_s$, equation (5.7.1) follows. Part (i) of theorem 5.7.3 is proved.

(ii) We continue to assume that E is compact, and we employ the stochastic representation constructed above. Its filtration is the basic filtration of $X_\bullet$. Let $\alpha_n > 0$ in $\mathbb{R}$ and $\gamma_n \ge 0$ in C be such that the sequence $\big(U_{\alpha_n}\gamma_n\big)$ is dense in C_{0+}. Let $Z^n_\bullet$ be the process

$$t \mapsto e^{-\alpha_n t} \cdot U_{\alpha_n}\gamma_n \circ X_t$$

of item 5.7.2 (ii). It is a global $L^2(\mathbb{P}^x)$-integrator for every single $x \in E$ (lemma 2.5.27). The set Osc of points $\omega \in \Omega$ where $\mathbb{Q} \ni q \mapsto Z^n_q(\omega)$ oscillates anywhere for any n belongs to $\mathcal{F}_\infty$ and is $\mathbb{P}$-nearly empty for every probability $\mathbb{P}$ with respect to which the Z^n are integrators (lemma 2.3.1), in particular for every one of the $\mathbb{P}^x$. Since the differences of the $U_{\alpha_n}\gamma_n$ are dense in C, we are assured that for all $\omega \in \Omega' \stackrel{\text{def}}{=} \Omega \setminus Osc$ and all $\phi \in C_0(E)$

$$\mathbb{Q} \ni q \mapsto \phi\big(X_q(\omega)\big) \text{ has left and right limits at all finite instants.}$$

Fix an instant t and an $\omega \in \Omega'$ and denote by $L_t(\omega)$ the intersection over n of the closures of the sets $\{X_q(\omega) : t \le q \le t + 1/n\} \subset E$. This set contains at least one point, by compactness, but not any more. Indeed, if there were two, we would choose a $\phi \in C$ that separates them; the path $q \mapsto \phi\big(X_q(\omega)\big)$ would have two limit points as $q \downarrow t$, which is impossible. Therefore

$$X'_t(\omega) \stackrel{\text{def}}{=} \lim_{\mathbb{Q} \ni q \downarrow t} X_q(\omega)$$

exists for every $t \ge 0$ and $\omega \in \Omega'$ and defines a right-continuous E-valued process $X'_\bullet$ on Ω'. A similar argument shows that $X'_\bullet$ has a unique left limit in E at all instants. The set $\big[X_t \ne X'_t\big]$ equals the union of the sets

$\big[Z^n_t \neq Z^n_{t+}\big] \in \mathcal{F}^0_{t+1}$, which are $\mathfrak{P}$-nearly empty in view of item 5.7.2 (i). Thus $X'_{.}$ is adapted to the $\mathfrak{P}$-regularization of $\mathcal{F}^0_{.}$. It is clearly also adapted to that filtration's right-continuous version $\mathcal{F}'_{.}$. Moreover, equation (5.7.1) stays when $\mathcal{F}^0_{.}$ is replaced by $\mathcal{F}'_{.}$. Namely, if $\mathbb{Q} \ni q_n \downarrow s$, then the left-hand side of

$$\mathbb{E}^x\big[\phi(X_t)|\mathcal{F}_{q_n}\big] = T_{t-q_n}\phi(X_{q_n})$$

converges in $\mathbb{E}^x$-mean to $\mathbb{E}^x[\phi(X_t)|\mathcal{F}'_s]$, while the right-hand side converges to $T_{t-s}\phi(X_s)$ by a slight extension of item 5.7.2 (i). Part (ii) is proved: the primed representation $\big(\Omega', \mathcal{F}'_{.}, X'_{.}, \mathfrak{P}\big)$ meets the description.

Let us then drop the primes and address the case of noncompact E. We let E^Δ denote the one-point compactification $E \cup \{\Delta\}$ of E, and consider on E^Δ the Feller semigroup $T^\Delta_{.}$ of remark A.9.6 on page 465. Functions in $C(E^\Delta)$ that vanish at Δ are identified with functions of $C_0(E)$. Let $\mathfrak{X}^\Delta \stackrel{\text{def}}{=} \big(\Omega, \mathcal{F}_{.}, X^\Delta_{.}, \mathfrak{P}^\Delta\big)$ be the corresponding stochastic representation provided by the proof above, with $\mathfrak{P}^\Delta = \{\mathbb{P}^x : x \in E^\Delta\}$. Note that $T^\Delta_t\big(x, \{\Delta\}\big) = 0$, which has the effect that $X^\Delta_t \neq \Delta$ $\mathbb{P}^x$-almost surely for all $x \in E$. Pick a function $\gamma^\Delta \in C(E^\Delta)$ that is strictly positive on E and vanishes at Δ and note that $U^\Delta_1\gamma^\Delta$ is of the same description. Then $t \mapsto Z^\Delta_t = e^{-t} \cdot U_1\gamma^\Delta \circ X^\Delta_t$ is a positive bounded right-continuous supermartingale on $\mathcal{F}_{.}$ with left limits and equals zero at time t if and only if $X^\Delta_t = \Delta$. Due to exercise 2.5.32, the path of Z^Δ is bounded away from zero on finite time-intervals, which means that $X^\Delta_{.}$ stays in a compact subset of E during any finite time-interval, except possibly in a $\mathfrak{P}$-nearly empty set N. Removing N from Ω leaves us with a stochastic representation of $T_{.}$ on E. It clearly satisfies (ii) of theorem 5.7.3.

Proof of Theorem 5.7.3 (iii). We start with the strong Markov property. It clearly suffices to prove equation (5.7.2) for the case $f \in C_0(E)$. The general case will follow from an application of the monotone class theorem. Let then $A \in \mathcal{F}_S$ and fix $\mathbb{P}^x \in \mathfrak{P}$. To start with, assume that S takes only countably many values. Then

$$\int A \cdot f\big(X_{S+\sigma}\big)\, d\mathbb{P}^x = \sum_{s \ge 0} \int [S = s] \cap A \cdot f\big(X_{S+\sigma}\big)\, d\mathbb{P}^x$$

since $[S = s] \cap A \in \mathcal{F}_s$:
$$= \sum_{s \ge 0} [S = s] \cap A \cdot \mathbb{E}^x\Big[f\big(X_{S+\sigma}\big)\big|\mathcal{F}_s\Big]\, d\mathbb{P}^x$$

by (5.7.1):
$$= \sum_{s \ge 0} \int [S = s] \cap A \cdot \mathbb{E}^{X_s}\big[f(X_\sigma)\big]\, d\mathbb{P}^x$$

as $X_s = X_S$ on $[S = s]$:
$$= \sum_{s \ge 0} \int [S = s] \cap A \cdot \mathbb{E}^{X_S}\big[f(X_\sigma)\big]\, d\mathbb{P}^x$$

$$= \int A \cdot \mathbb{E}^{X_S}\Big[f\big(X_\sigma\big)\Big]\, d\mathbb{P}^x\ .$$

In the general case let $S^{(n)}$ be the approximating stopping times of exercise 1.3.20. Since $x \mapsto \mathbb{E}^x\big[f(X_\sigma)\big]$ is continuous and bounded, taking the limit in

$$\int A \cdot f\big(X_{S^{(n)}+\sigma}\big)\, d\mathbb{P}^x = \int A \cdot \mathbb{E}^{X_{S^{(n)}}}\Big[f\big(X_\sigma\big)\Big]\, d\mathbb{P}^x$$

produces the desired equality

$$\int A \cdot f\big(X_{S+\sigma}\big)\, d\mathbb{P}^x = \int A \cdot \mathbb{E}^{X_S}\Big[f\big(X_\sigma\big)\Big]\, d\mathbb{P}^x .$$

This proves the strong Markov property.

Now to the quasi-left-continuity. Let $\phi, \psi \in C_0(E)$. Thanks to the right-continuity of the paths

$$X_T = \lim_{t\downarrow 0} \lim_{n\to\infty} X_{T_n+t} ,$$

and therefore

$$\begin{aligned}
\mathbb{E}^x\Big[\phi\big(X_{T-}\big)\cdot\psi\big(X_T\big)\Big] &= \lim_{t\downarrow 0}\lim_{n\to\infty} \mathbb{E}^x\Big[\phi\big(X_{T_n}\big)\cdot\psi\big(X_{T_n+t}\big)\Big] \\
&= \lim_{t\downarrow 0}\lim_{n\to\infty} \mathbb{E}^x\Big[\phi\big(X_{T_n}\big)\cdot\big(T_t\psi\big)\big(X_{T_n}\big)\Big] \\
&= \lim_{t\downarrow 0} \mathbb{E}^x\Big[\phi\big(X_{T-}\big)\cdot\big(T_t\psi\big)\big(X_{T-}\big)\Big] \\
&= \mathbb{E}^x\Big[\phi\big(X_{T-}\big)\cdot\psi\big(X_{T-}\big)\Big] .
\end{aligned}$$

The equality

$$\mathbb{E}^x\Big[f(X_{T-}, X_T)\Big] = \mathbb{E}^x\Big[f(X_{T-}, X_{T-})\Big]$$

therefore holds for functions on $E \times E$ of the form $(x,y) \mapsto \sum_i \phi_i(x)\psi_i(y)$, which form a multiplicative class generating the Borels of $E \times E$. Thus $\mathbb{E}^x\big[h(X_{T-} - X_T)\big] = 0$ for all Borel functions h on E, which implies that $\big[X_{T-} \neq X_T\big] = \bigcup_n \big[X_{(T\wedge n)-} \neq X_{T\wedge n}\big]$ is $\mathbb{P}^x$-nearly empty: the quasi-left-continuity follows.

Finally let us address the regularity. Let us denote by $\mathcal{X}^0_{\bullet} \subseteq \mathcal{F}_{\bullet}$ the basic filtration of $X_{\bullet}$. Fix a $\mathbb{P}^x \in \mathfrak{P}$, a $t \geq 0$, and a bounded $\mathcal{X}^0_{t+}$-measurable function F. Set $G \stackrel{\text{def}}{=} F - \mathbb{E}^x[F|\mathcal{X}^0_t]$. This function is measurable on $\mathcal{X}^0_\tau$ for all $\tau > t$. Now let $f : \mathscr{D}_b(E) \to \mathbb{R}$ be a bounded function of the form

$$f(\omega) = f^t(\omega)\cdot\phi\big(X_u(\omega)\big) , \tag{$*$}$$

where $f^t \in \mathcal{X}^0_t$, $\phi \in C_0(E)$, and $u > t$, and consider

$$I^f \stackrel{\text{def}}{=} \int f \cdot G\, d\mathbb{P}^x = \int f^t \cdot G \cdot \phi(X_u)\, d\mathbb{P}^x . \tag{$**$}$$

Pick a $\tau \in (t,u)$ and condition on $\mathcal{X}^0_\tau$. Then

$$I^f = \int f^t \cdot G \cdot T_{u-\tau}\phi(X_\tau)\, d\mathbb{P}^x \,.$$

Due to item 5.7.2 (i), taking $\tau \downarrow t$ produces

$$I^f = \int f^t \cdot G \cdot T_{u-t}\phi(X_t)\, d\mathbb{P}^x \,.$$

The factor of G in this integral is measurable on $\mathcal{X}^0_t$, so the very definition of G results in $I^f = 0$.

Now let $\overline{f}(\omega) = \overline{f}^t(\omega) \cdot \overline{\phi}\big(X_v(\omega)\big)$ be a second function of the form $(*)$, where $v \ge u$, say. Conditioning on $\mathcal{X}^0_u$ produces

$$I^{f\overline{f}} = \int f^t\overline{f}^t \cdot G \cdot \phi(X_u)\overline{\phi}(X_v)\, d\mathbb{P}^x = \int f^t\overline{f}^t \cdot G \cdot \big(\phi \cdot T_{v-u}\overline{\phi}\big)(X_u)\, d\mathbb{P}^x \,.$$

This is an integral of the form $(**)$ and therefore vanishes. That is to say, $\int f \cdot G\, d\mathbb{P}^x{=}0$ for all f in the algebra generated by functions of the form $(*)$. Now this algebra generates $\mathcal{X}^0_\infty$. We conlude that $[G \ne 0]$ is $\mathbb{P}^x$-negligible. Since this set belongs to $\mathcal{X}^0_{t+1}$, it is even $\mathbb{P}^x$-nearly empty, and thus $\mathbb{P}^x$-nearly $F = \mathbb{E}^x[F|\mathcal{X}^0_t]$. To summarize: if $F \in \mathcal{X}^0_{t+}$, then F differs $\mathbb{P}^x$-nearly from some $\mathcal{X}^0_t$-measurable function $F^{\mathbb{P}^x} = \mathbb{E}^x[F|\mathcal{X}^0_t]$, and this holds for all $\mathbb{P}^x \in \mathfrak{P}$. This says that F belongs to the regularization $\mathcal{X}^{\mathfrak{P}}_t$. Thus $\mathcal{X}^0_{t+} \subset \mathcal{X}^{\mathfrak{P}}_t$ and then $\mathcal{X}^{\mathfrak{P}}_{t+} = \mathcal{X}^{\mathfrak{P}}_t \quad \forall\, t \ge 0$. Theorem 5.7.3 is proved in its entirety. ∎

Exercise 5.7.7 Let $\mathfrak{X}^{\vdash} = \big(\Omega, \mathcal{F}_., X^{\vdash}_., \{\mathbb{P}^x\}\big)$ be a regular stochastic representation of $T^{\vdash}_.$. Describe $\mathfrak{X}^{\vdash}$ and the projection $X_. \stackrel{\text{def}}{=} \pi_E \circ X^{\vdash}_.$ on the second factor.

Exercise 5.7.8 To appreciate the designation "Zero-One Law" of theorem 5.7.3 (iiic) show the following: for any $x \in E$ and $A \in \mathcal{F}^{\mathfrak{P}}_{0+}[X_.]$, $\mathbb{P}^x[A]$ is either zero or one. In particular, for $B \in \mathcal{B}^\bullet(E)$,

$$T^B \stackrel{\text{def}}{=} \inf\{t > 0 : X_t \in B\}$$

is $\mathbb{P}^x$-almost surely either strictly positive or identically zero.

Exercise 5.7.9 (The Canonical Representation of $T_.$) Let $\mathfrak{X} = \big(\Omega, \mathcal{F}_., X_., \{\mathbb{P}^x\}\big)$ be a regular stochastic representation of the conservative Feller semigroup $T_.$. It gives rise to a map $\rho : \Omega \to \mathscr{D}_E$, space of right-continuous paths $x_. : [0,\infty) \to E$ with left limits, via $\rho(\omega)_t = X_t(\omega)$. Equip $\mathscr{D}_E$ with its basic filtration $\mathcal{F}^0_.[\mathscr{D}_E]$, the filtration generated by the evaluations $x_. \mapsto x_t$, $t \in [0,\infty)$, which we denote again by X_t. Then ρ is $\mathcal{F}^0_\infty[\mathscr{D}_E]/\mathcal{F}_\infty$-measurable, and we may define ***the laws of*** $\mathfrak{X}$ as the images under ρ of the $\mathbb{P}^x$ and denote them again by $\mathbb{P}^x$. They depend only on the semigroup $T_.$, not on its representation $\mathfrak{X}$. We now replace $\mathcal{F}^0_.[\mathscr{D}_E]$ on $\mathscr{D}_E$ by the natural enlargement $\mathcal{F}^{\mathfrak{P}}_{.+}[\mathscr{D}_E]$, where $\mathfrak{P} = \{\mathbb{P}^x : x \in E\}$, and then rename $\mathcal{F}^{\mathfrak{P}}_{.+}[\mathscr{D}_E]$ to $\mathcal{F}_.$. The regular stochastic representation $\big(\mathscr{D}_E, \mathcal{F}_., X_., \{\mathbb{P}^x\}\big)$ is the canonical representation of $T_.$.

Exercise 5.7.10 (Continuation) Let us denote the typical path in $\mathscr{D}_E$ by ω, and let $\theta_s : \mathscr{D}_b(E) \to \mathscr{D}_b(E)$ be the ***time shift operator*** on paths defined by

$$(\theta_s(\omega))_t = \omega_{s+t} \ , \qquad\qquad s,t \geq 0 \ .$$

Then $\theta_s \circ \theta_t = \theta_{s+t}$, $X_t \circ \theta_s = X_{t+s}$, $\theta_s \in \mathcal{F}_{s+t}/\mathcal{F}_t$, and $\theta_s \in \mathcal{F}^{\mathfrak{P}}_{s+t}/\mathcal{F}^{\mathfrak{P}}_t$ for all $s,t \geq 0$, and for any finite $\mathcal{F}$-stopping time S and bounded $\mathcal{F}$-measurable random variable F

$$\mathbb{E}^x\big[F \circ \theta_S | \mathcal{F}_S\big] = \mathbb{E}^{X_S}[F] \ :$$

"the semigroup $T_{\boldsymbol{.}}$ is represented by the flow $\theta_{\boldsymbol{.}}$."

Exercise 5.7.11 Let E be $\mathbb{N}$ equipped with the discrete topology and define the ***Poisson semigroup*** T_t by

$$(T_t\phi)(k) = e^{-t}\sum_{i=0}^{\infty}\phi(k+i)\frac{t^i}{i!} \ , \qquad\qquad \phi \in C_0(\mathbb{N}) \ .$$

This is a Feller semigroup whose generator $A\phi : n \mapsto \phi(n+1) - \phi(n)$ is defined for all $\phi \in C_0(\mathbb{N})$. Any regular process representing this semigroup is ***Poisson process***.

Exercise 5.7.12 Fix a $t > 0$, and consider a bounded continuous function defined on *all* bounded paths $\omega : [0,\infty) \to E$ that is continuous in the topology of pointwise convergence of paths and depends on the path prior to t only; that is to say, if the stopped paths ω^t and ω'^t agree, then $F(\omega) = F(\omega')$. (i) There exists a countable set $\tau \in [0,t]$ such that F is a cylinder function based on τ; in other words, there is a function f defined on all bounded paths $\xi : \tau \to E$ and continuous in the product topology of E_τ such that $F = f(X_\tau)$. (ii) The function $x \mapsto \mathbb{E}^x[F]$ is continuous. (iii) Let $T_{n,\boldsymbol{.}}$ be a sequence of Feller semigroups converging to $T_{\boldsymbol{.}}$ in the sense that $T_{n,t}\phi(x) \to T_t\phi(x)$ for all $\phi \in C_0(E)$ and all $x \in E$. Then $\mathbb{E}^x_n[F] \to \mathbb{E}^x[F]$.

Exercise 5.7.13 For every $x \in E$, $t > 0$, and $\epsilon > 0$ there exist a compact set K such that

$$\mathbb{P}^x\big[X_s \in K \quad \forall\, s \in [0,t]\big] > 1 - \epsilon \ . \tag{5.7.6}$$

Exercise 5.7.14 Assume the semigroup $T_{\boldsymbol{.}}$ is compact; that is to say, the image under T_t of the unit ball of $C_0(E)$ is compact, for arbitrarily small, and then all, $t > 0$. Then T_t maps bounded Borel functions to continuous functions and $x \mapsto \mathbb{E}^x[F\circ\theta_t]$ is continuous for bounded $F \in \mathcal{F}_\infty$, provided $t > 0$. Equation (5.7.6) holds for all x in an open set.

Theorem 5.7.15 (A Feynman–Kac Formula) *Let $T_{s,t}$ be a conservative family of Feller transition probabilities, with corresponding infinitesimal generators A_t, and let $\mathfrak{X}^{\vdash} = \big(\Omega, \mathcal{F}_{\boldsymbol{.}}, X^{\vdash}_{\boldsymbol{.}}, \{\mathbb{P}^x\}\big)$ be a regular stochastic representation of its time-rectification $T^{\vdash}_{\boldsymbol{.}}$. Denote by $X_{\boldsymbol{.}}$ its trace on E, so that $X^{\vdash}_t = (t, X_t)$.*

Suppose that $\Phi \in \operatorname{dom}(\breve{A}^{\vdash})$ satisfies on $[0,u] \times E$ the ***backward equation***

$$\frac{\partial\Phi}{\partial t}(t,x) + A_t\Phi\,(t,x) = \big(q{\cdot}\Phi - g\big)(t,x) \tag{5.7.7}$$

and the final condition $\qquad \Phi(u,x) = f(x)$, *where $q, g : [0,u] \times E \to \mathbb{R}$ and $f : E \to \mathbb{R}$ are continuous. Then $\Phi(t,x)$ has the following stochastic representation:*

$$\Phi(t,x) = \mathbb{E}^{t,x}\big[Q_u f(X_u)\big] + \mathbb{E}^{t,x}\Big[\int_t^u Q_\tau g(X^{\vdash}_\tau)\, d\tau\Big] , \qquad (5.7.8)$$

where $$Q_\tau \stackrel{\text{def}}{=} \exp\Big(-\int_t^\tau q(X^{\vdash}_s)\, ds\Big) ,$$

provided *(a) q is bounded below, (b) $g \in \check{C}$ or $g \geq 0$, (c) $f \in \check{C}$ or $f \geq 0$,* ***and***

(d) $X^{\vdash}_{\bullet}$ *has continuous paths* *or*

(d') $\int_{E^{\vdash}} |\Phi(s,y)|^p \, T^{\vdash}_t\big((x,t), ds \times dy\big)$ *is finite for some $p > 1$.*

Proof. Let $S \leq u$ be a stopping time and set $G_v = \int_t^v Q_\tau g(X^{\vdash}_\tau)\, d\tau$. Itô's formula gives $\mathbb{P}^{t,x}$-almost surely

$$\begin{aligned}
G_S + Q_S\Phi(X^{\vdash}_S) - \Phi(t,x) &= G_S + Q_S\Phi(X^{\vdash}_S) - Q_t\Phi(X^{\vdash}_t) \\
&= G_S + \int_{t+}^S \Phi(X^{\vdash}_\tau)\, dQ_\tau + \int_{t+}^S Q_\tau \, d\Phi(X^{\vdash}_\tau) \\
&= G_S - \int_t^S Q_\tau \cdot q\Phi{\circ}X^{\vdash}_\tau \, d\tau \\
\text{by 5.7.6:} \qquad & \qquad + \int_t^S Q_\tau \cdot \breve{A}^{\vdash}\Phi{\circ}X^{\vdash}_\tau \, d\tau + \int_{t+}^S Q_\tau \, dM^\Phi_\tau \\
&= \int_t^S Q_\tau \cdot \big(g - q\Phi\big){\circ}X^{\vdash}_\tau \, d\tau \\
\text{by A.9.15 and (5.7.7):} \qquad & \qquad + \int_t^S Q_\tau \cdot (q\Phi - g){\circ}X^{\vdash}_\tau \, d\tau + \int_{t+}^S Q_\tau \, dM^\Phi_\tau \\
&= \int_{t+}^S Q_\tau \, dM^\Phi_\tau \ .
\end{aligned}$$

Since the paths of $X^{\vdash}_{\bullet}$ stay in compact sets $\mathbb{P}^{t,x}$-almost surely and all functions appearing are continuous, the maximal function of every integrand above is $\mathbb{P}^{t,x}$-almost surely finite at time S, in the $d\Phi(X^{\vdash}_\tau)$- and dM^Φ_τ-integrals. Thus every integral makes sense (theorem 3.7.17 on page 137) and the computation is kosher. Therefore

$$\begin{aligned}
\Phi(t,x) &= Q_S\Phi(S, X_S) + \int_t^S Q_\tau g(X^{\vdash}_\tau)\, d\tau - \int_{t+}^S Q_\tau \, dM^\Phi_\tau \\
&= \mathbb{E}^{t,x}\big[Q_S\Phi(S,X_S)\big] + \mathbb{E}^{t,x}\Big[\int_t^S Q_\tau g(X^{\vdash}_\tau)\, d\tau\Big] - \mathbb{E}^{t,x}\Big[\int_{t+}^S Q_\tau \, dM^\Phi_\tau\Big] ,
\end{aligned}$$

provided the random variables have finite $\mathbb{P}^{t,x}$-expectation. The proviso in the statement of the theorem is designed to achieve this and to have the last expectation vanish. The assumption that q be bounded below has the effect that $Q^\star_u$ is bounded. If $g \geq 0$, then the second expectation exists at time u.

The desired equality equation (5.7.8) now follows upon application of $\mathbb{E}^{t,x}$, and it is to make this expectation applicable that assumptions (a)–(d) are needed. Namely, since q is bounded below, Q is bounded above. The solidity of $\breve{C}$ together with assumptions (b) and (c) make sure that the expectation of the first two integrals exists. If (d') is satisfied, then $M^\Phi_{\bullet}$ is an L^1-integrator (theorem 2.5.30 on page 85) and the expectation of $\int_t^u Q_\tau \, dM^\Phi_\tau$ vanishes. If (d) is satisfied, then $X \in K_{n+1}$ up to and incuding time S_n, so that M^Φ stopped at time S_n is a bounded martingale: we take the expectation at time S_n, getting zero for the martingale integral, and then let $n \to \infty$. ∎

Repeated Footnotes: 271[1] 272[2] 273[3] 274[4] 277[5] 278[7] 280[8] 281[10] 282[11] 282[12] 287[16] 288[17] 293[20] 295[21] 297[23] 301[25] 303[26] 303[28] 305[30] 308[32] 310[33] 312[35] 312[36] 319[37] 320[38] 321[39] 323[40] 334[44] 352[48] 354[50]