April 25, 2013 (last update: March 14, 2016)

# LECTURE NOTES ON THE YAMADA–WATANABE CONDITION FOR THE PATHWISE UNIQUENESS OF SOLUTIONS OF CERTAIN STOCHASTIC DIFFERENTIAL EQUATIONS

SÜHAN ALTAY AND UWE SCHMOCK

Abstract. In these lecture notes we discuss the Yamada–Watanabe condition for the pathwise uniqueness of the solution of certain stochastic differential equations. This condition is weaker than the usual Lipschitz condition, the proof is based on Bihari's inequality. An important application in mathematical finance is the Cox–Ingersoll–Ross model.

Contents



## 1. Introduction and basic definitions

These lecture notes explain the Yamada–Watanabe condition, which relaxes the Lipschitz condition for the pathwise uniqueness of solutions of stochastic differential equations (SDEs) of the type

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dW_t\,, \quad t \ge 0.$$

Hence, this condition can be used to show the strong uniqueness of solutions of a SDE with certain non-Lipschitz coefficients. The main references for this part are [3], [4], [9] and [10].

In mathematical finance, this is of particular interest for the Cox–Ingersoll–Ross model (CIR model for short), which describes the stochastic evolution of interest rates $(r_t)_{t\ge 0}$ by the SDE

$$dr_t = \alpha(\mu - r_t)\,dt + \sigma\sqrt{r_t}\,dW_t, \quad t \ge 0,$$

with $r_0 \ge 0$, where $\alpha$, $\mu$ with $\alpha\mu \ge 0$ and $\sigma$ denote real constants. The Yamada–Watanabe condition also gives the strong uniqueness of the solution of the SDE defining the squared Bessel process of 'dimension' $\delta \in [0, \infty)$, cf. [7, Ch. XI, §1, Def. (1.1)], which can be used to express the interest rate process in the CIR model.

The work on these lecture notes was financially supported by the Christian Doppler Research Association (CDG). The authors gratefully acknowledge the fruitful collaboration and support by the Bank Austria and the Austrian Federal Financing Agency (ÖBFA) through CDG and the CD-Laboratory for Portfolio Risk Management (PRisMa Lab) http://www.prismalab.at/.

Before stating the main theorem, we start with some definitions necessary for the sequel.

**Definition 1.1.** Given two jointly Borel measurable functions $b\colon [0,\infty)\times\mathbb{R}^n \to \mathbb{R}^n$ and $\sigma\colon [0,\infty)\times\mathbb{R}^n \to \mathbb{R}^{n\times d}$ and a probability measure $\mu$ on $(\mathbb{R}^n,\mathcal{B}_n)$, a solution of the stochastic differential equation

$$dX_t = b(t,X_t)\,dt + \sigma(t,X_t)\,dW_t, \quad t\ge 0, \tag{1.2}$$

with initial distribution $\mu$ is a pair $(W,X)$ of continuous adapted processes defined on a filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ with $\mathbb{F}=(\mathcal{F}_t)_{t\ge0}$ such that

(a) $W=(W_t)_{t\ge0}$ is a standard $(\mathbb{F},\mathbb{P})$-Brownian motion with values in $\mathbb{R}^d$,
(b) the initial value $X_0$ has distribution $\mu$,
(c) the integrals implicitly given by (1.2) are well defined, i.e., for all $t\ge 0$, $i\in\{1,\dots,n\}$ and $j\in\{1,\dots,d\}$, the corresponding component functions of $b$ and $\sigma$ satisfy

$$\int_0^t \sigma_{ij}^2(s,X_s)\,ds \overset{\text{a.s.}}{<} \infty \qquad\text{and}\qquad \int_0^t |b_i(s,X_s)|\,ds \overset{\text{a.s.}}{<} \infty,$$

(d) for every $i\in\{1,2,\dots,n\}$, the $i$-th component process of

$$X = \bigl((X_t^{(1)},\dots,X_t^{(n)})^\top\bigr)_{t\ge0}$$

satisfies, up to indistinguishability,

$$X_t^{(i)} = X_0^{(i)} + \int_0^t b_i(s,X_s)\,ds + \sum_{j=1}^d \int_0^t \sigma_{ij}(s,X_s)\,dW_s^j, \quad t\ge 0. \tag{1.3}$$

**Definition 1.4.** We say that there is pathwise uniqueness for the SDE (1.2) with initial distribution $\mu$, if whenever $(W,X)$ and $(\tilde W,\tilde X)$ are two solutions of (1.2) defined on the same filtered probability space with $W=\tilde W$ (same Brownian motion) and $X_0 \overset{\text{a.s.}}{=} \tilde X_0$ (same $\mathcal{F}_0$-measurable initial condition with distribution $\mu$), then $X$ and $\tilde X$ are indistinguishable, that is, there exists a set $N\in\sigma(\bigcup_{t\ge0}\mathcal{F}_t)$ with $\mathbb{P}[N]=0$ such that $\{X_t\ne\tilde X_t\}\subset N$ for all $t\in[0,\infty)$.

The next example shows that pathwise uniqueness may depend on the initial distribution $\mu$.

**Example 1.5.** Fix $\alpha\in(0,1)$ and consider the (deterministic) one-dimensional SDE

$$dX_t = \frac{(X_t^+)^\alpha}{1-\alpha}\,dt, \qquad t\ge 0,$$

where $x^+ := \max\{x,0\}$ denotes the positive part of $x\in\mathbb{R}$. If $\mu$ assigns probability one to the interval $(0,\infty)$, then $X_t := (X_0^{1-\alpha}+t)^{1/(1-\alpha)}$ for $t\ge0$ with $\mathcal{L}(X_0)=\mu$ is the unique solution (up to a $\mathbb{P}$-null set) as defined above, because for every $\omega\in\Omega$ with $X_0(\omega)>0$, the solution has non-negative derivative everywhere and therefore stays inside the interval $[X_0(\omega),\infty)$ and the $\alpha$-power function is Lipschitz continuous in $(\frac12 X_0(\omega),\infty)$, hence standard uniqueness results for initial value problems apply, c.f. [8, Theorem 2.5]. If $\mu$ is the Dirac measure in 0, then for every $\mathbb{F}$-stopping time $\tau\colon\Omega\to[0,\infty)$, the process

$$X_t := \bigl(\max\{t-\tau,0\}\bigr)^{1/(1-\alpha)}, \qquad t\ge0,$$

is a solution.

## 2. Yamada–Watanabe condition for SDEs

In this section, we give the main result, which combines the one- and multi-dimensional case. However, we mention that the Yamada–Watanabe condition is essentially a one-dimensional result (see Remarks 2 and 3 in [10]). For the one-dimension setting, there is also an approach to pathwise uniqueness using local times, see [7, Ch. IX, §3].

The following theorem is the main result of these lectures notes; for its proof we assume that the filtration is right-continuous. We use $|\cdot|$ for the $n$-dimensional Euclidean norm and $\|\cdot\|_{\mathrm{F}}$ for the Frobenius matrix norm, see Remark 2.20 below.

**Theorem 2.1.** *Consider the stochastic differential equation* (1.2)*. Assume that there exist a constant $\gamma > 0$ and functions $\kappa, \varrho\colon [0,\gamma] \to [0,\infty)$ satisfying $\kappa(0) = 0$,*

$$|b(t,x) - b(t,y)| \le \kappa(|x-y|), \tag{2.2}$$

*and*

$$\|\sigma(t,x) - \sigma(t,y)\|_{\mathrm{F}} \le \varrho(|x-y|) \tag{2.3}$$

*for all $t \in [0,\infty)$ and $x$, $y \in \mathbb{R}^n$ with $|x-y| \le \gamma$. Furthermore, assume that $\varrho$ is non-decreasing, $\varrho(u) > 0$ for all $u \in (0,\gamma]$ and its square satisfies the Osgood condition*[1]*, i.e.,*

$$\int_0^\gamma \frac{du}{\varrho^2(u)} = \infty. \tag{2.4}$$

*In addition, assume that there exists a non-decreasing, concave and continuous function $G\colon [0,\gamma] \to [0,\infty)$ with $G(0) = 0$, strictly positive on $(0,\gamma]$, such that*

$$G(u) \ge \kappa(u) + \frac{n-1}{2u}\varrho^2(u) \qquad \forall\, u \in (0,\gamma] \tag{2.5}$$

*and it also satisfies the Osgood condition*

$$\int_0^\gamma \frac{du}{G(u)} = \infty. \tag{2.6}$$

*Then the pathwise uniqueness of solutions of* (1.2) *holds for every initial distribution $\mu$.*

*Remark* 2.7. Note that for $n = 1$ with the choice $G(u) = \kappa(u)$ for $u \in [0,\gamma]$, the conditions on $G$ are actually conditions on $\kappa$, in particular (2.6) reduces to the Osgood condition

$$\int_0^\gamma \frac{du}{\kappa(u)} = \infty. \tag{2.8}$$

For $n \ge 2$ and vanishing drift $b$, we can choose $\kappa$ to be the zero function. With the choice $G(u) = \frac{n-1}{2u}\varrho^2(u)$ for $u \in [0,\gamma]$, the condition (2.6) is equivalent to

$$\int_0^\gamma \frac{u}{\varrho^2(u)}\, du = \infty, \tag{2.9}$$

which is substantially more restrictive than (2.4).

**Example 2.10.** For the continuous concave function $\kappa$ satisfying the Osgood condition (2.8), typical examples are $\kappa(u) = Cu$ with a constant $C > 0$ as well as

$$\kappa(u) = \begin{cases} 0 & \text{for } u = 0, \\ Cu \log \frac{1}{u} & \text{for } u \in (0,\gamma], \end{cases} \tag{2.11}$$

[1] Named after William Fogg Osgood, cf. [5, p. 344]

with $\gamma := e^{-2}$, which has unbounded slope close to the origin. The example (2.11) satisfies (2.8), because for $\delta \in (0, \gamma]$

$$\int_\delta^\gamma \frac{du}{u \log \frac{1}{u}} = \log\log\frac{1}{\delta} - \log\log\frac{1}{\gamma} \to \infty \quad \text{as } \delta \searrow 0.$$

For the function $\varrho$ satisfying (2.4), one can always take $\varrho(u) = \sqrt{\kappa(u)}$ for $u \in [0, \gamma]$, where $\kappa$ satisfies (2.8), and also $\varrho(u) = Cu^\alpha$ with exponent $\alpha \geq 1/2$ is possible. If $\kappa$ satisfies (2.8), then $\varrho(u) := \sqrt{u\kappa(u)}$ for $u \in [0, \gamma]$ satisfies (2.9) and $G(u) := \kappa(u) + \frac{n-1}{2u}\varrho^2(u)$ for $u \in [0, \gamma]$ satisfies (2.6) for every $n \in \mathbb{N}$. Note that the Lipschitz-continuous case corresponding to $\varrho(u) = \kappa(u) = Cu$ is included.

**Example 2.12** (Extension of (2.11))**.** Define $\exp^{\circ 0}(u) = u$ and $\log^{\circ 0}(u) = u$ for all $u \in \mathbb{R}$. Then, for all $l \in \mathbb{N}$, define iteratively the $l$-fold composition of the exponential function by $\exp^{\circ l}(u) = \exp(\exp^{\circ l-1}(u))$ for $u \in \mathbb{R}$ and the $l$-fold composition of the natural logarithm by $\log^{\circ l}(u) = \log(\log^{\circ l-1}(u))$ for all $u > \exp^{\circ l-1}(0)$. With this notation and a constant $C > 0$ define for $l \in \mathbb{N}_0$ the constant $\gamma_l = 1/\exp^{\circ l}(2) > 0$ and the function

$$\kappa_l(u) = \begin{cases} 0 & \text{for } u = 0, \\ Cu \prod_{k=1}^l \log^{\circ k}\left(\frac{1}{u}\right) & \text{for } u \in (0, \gamma_l]. \end{cases} \tag{2.13}$$

Note that for every $l \in \mathbb{N}_0$

$$\frac{d}{du} \log^{\circ l+1}\left(\frac{1}{u}\right) = -\frac{C}{\kappa_l(u)}, \quad u \in (0, \gamma_l]. \tag{2.14}$$

The example (2.13) satisfies the Osgood condition (2.8), because by (2.14) and the fundamental theorem of calculus, for $\delta \in (0, \gamma_l]$,

$$C \int_\delta^{\gamma_l} \frac{du}{\kappa_l(u)} = \log^{\circ l+1}\left(\frac{1}{\delta}\right) - \log^{\circ l+1}\left(\frac{1}{\gamma_l}\right) \to \infty \quad \text{as } \delta \searrow 0.$$

It follows from (2.13) and (2.14) using the product rule that, for every $l \in \mathbb{N}$,

$$\kappa_l(u) = \kappa_{l-1}(u) \log^{\circ l}\left(\frac{1}{u}\right),$$

$$\kappa_l'(u) = \kappa_{l-1}'(u) \log^{\circ l}\left(\frac{1}{u}\right) - C, \tag{2.15}$$

$$\kappa_l''(u) = \kappa_{l-1}''(u) \log^{\circ l}\left(\frac{1}{u}\right) - C\frac{\kappa_{l-1}'(u)}{\kappa_{l-1}(u)} \tag{2.16}$$

for all $u \in (0, \gamma_l]$. Since $\kappa_0'(u) \equiv C$, $\kappa_0''(u) \equiv 0$, and $\log^{\circ l}\left(\frac{1}{u}\right) \geq 2$ for all $u \in (0, \gamma_l]$ and $l \in \mathbb{N}$, it follows from (2.15) and (2.16) by induction that $\kappa_l'(u) \geq C$ and $\kappa_l''(u) \leq 0$ for all $u \in (0, \gamma_l]$ and $l \in \mathbb{N}_0$. Therefore, $\kappa_l$ is strictly increasing and concave on $[0, \gamma_l]$ for every $l \in \mathbb{N}_0$.

*Remark* 2.17. We will show below, contrary to the preceding example, that for every $\varepsilon > 0$ and $l \in \mathbb{N}$, the function $\tilde{\kappa}_{\varepsilon,l}\colon [0, \gamma_l] \to [0, \infty)$ with

$$\tilde{\kappa}_{\varepsilon,l}(u) := \begin{cases} 0 & \text{for } u = 0, \\ \kappa_{l-1}(u)\left(\log^{\circ l}\left(\frac{1}{u}\right)\right)^{1+\varepsilon} & \text{for } u \in (0, \gamma_l], \end{cases}$$

is growing too fast at the origin to satisfy the Osgood condition (2.8). Indeed, it follows using (2.14) that

$$\frac{d}{du}\Bigl(\log^{\circ l}\Bigl(\frac{1}{u}\Bigr)\Bigr)^{-\varepsilon} = -\frac{\varepsilon}{\bigl(\log^{\circ l}\bigl(\frac{1}{u}\bigr)\bigr)^{1+\varepsilon}}\frac{d}{du}\log^{\circ l}\Bigl(\frac{1}{u}\Bigr) = \frac{C\varepsilon}{\tilde{\kappa}_{\varepsilon,l}(u)}, \quad u \in (0,\gamma_l],$$

hence we have for every $\delta \in (0,\gamma_l]$ that

$$C\varepsilon\int_\delta^{\gamma_l}\frac{du}{\tilde{\kappa}_{\varepsilon,l}(u)} = \frac{1}{2^\varepsilon} - \frac{1}{\bigl(\log^{\circ l}\bigl(\frac{1}{\delta}\bigr)\bigr)^{\varepsilon}} \nearrow \frac{1}{2^\varepsilon} \quad \text{as } \delta \searrow 0.$$

*Remark* 2.18 (Deterministic dynamics). By setting the diffusion function $\sigma$ equal to zero and $\varrho(u) = \sqrt{u\kappa(u)}$ for $u \in [0,\infty)$, Theorem 2.1 also weakens the Lipschitz condition for the uniqueness of solutions of ordinary initial value problems to (2.8). Note that the function $\kappa$ (or $G$, respectively) doesn't need to be concave in this case, because the application of Jensen's inequality for (2.40) is not necessary (and the entire proof simplifies substantially).

**Corollary 2.19** (Application to the CIR model). *In one dimension, if the drift $b$ is Lipschitz continuous, and $\sigma$ is Hölder continuous with exponent* 1/2*, i.e., there exists some constant $C > 0$ such that*

$$|\sigma(t,x) - \sigma(t,y)| \le C\sqrt{|x-y|} \qquad \forall\ 0 \le t < \infty \ \text{ and } x,\, y \in \mathbb{R},$$

*then the pathwise uniqueness for the solution of equation* (1.2) *holds for every initial distribution $\mu$.*

*Remark* 2.20 (Review of linear algebra).

(a) On the vector space $\mathbb{R}^{n\times d}$ of $(n\times d)$-dimensional matrices, an inner product is defined by $\langle A, B\rangle = \operatorname{tr}(AB^\top)$ for all $A, B \in \mathbb{R}^{n\times d}$, there $\operatorname{tr}(AB^\top)$ denotes the trace of the $(n \times n)$-dimensional matrix $AB^\top$, i.e. the sum of all its entries on the diagonal,

$$\operatorname{tr}(AB^\top) = \sum_{i=1}^n (AB^\top)_{i,i} = \sum_{i=1}^n\sum_{j=1}^d A_{i,j}B_{i,j} = \operatorname{tr}(B^\top A), \tag{2.21}$$

where the last equality indicates that the $(d \times d)$-dimensional matrix $B^\top A$ has the same trace as $AB^\top$. Note that $\langle A, B\rangle$ equals the usual inner product on $\mathbb{R}^{nd}$, when the rectangular form of the matrices $A, B$ is ignored and they are viewed as vectors in $\mathbb{R}^{nd}$. The corresponding norm $\|\cdot\|_{\mathrm{F}}$ is the Frobenius matrix norm, given by $\|A\|_{\mathrm{F}} = \sqrt{\operatorname{tr}[AA^\top]}$. By the Cauchy–Schwarz inequality,

$$\bigl|\operatorname{tr}(AB^\top)\bigr| \le \sqrt{\operatorname{tr}[AA^\top]}\sqrt{\operatorname{tr}[BB^\top]}, \quad A, B \in \mathbb{R}^{n\times d}. \tag{2.22}$$

(b) Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix, i.e. $A = A^\top$. Then there exists an orthogonal matrix $U \in \mathbb{R}^{n\times n}$, whose columns form an orthonormal basis of eigenvectors of $A$ and which therefore satisfies $U^\top U = I$, and a corresponding diagonal matrix $D \in \mathbb{R}^{n\times n}$ with the real eigenvalues $\lambda_1, \dots, \lambda_n$ of $A$ on the diagonal such that $A = UDU^\top$. For every $r \in \mathbb{N}_0$ we can define $D^r \in \mathbb{R}^{n\times n}$ as the diagonal matrix with $\lambda_1^r, \dots, \lambda_n^r$ on the diagonal (using the convention $0^0 := 1$) and the matrix power $A^r = UD^rU^\top$, which is again symmetric. For $r, s \in \mathbb{N}_0$ we have the rule

$$A^rA^s = UD^rU^\top UD^sU^\top = UD^{r+s}U^\top = A^{r+s}. \tag{2.23}$$

Using (2.21), it follows that

$$\operatorname{tr}(A^r) = \operatorname{tr}(UD^rU^\top) = \operatorname{tr}(U^\top UD^r) = \operatorname{tr}(D^r) = \lambda_1^r + \cdots + \lambda_n^r. \tag{2.24}$$

If $A \in \mathbb{R}^{n\times n}$ is positive semi-definite, then the eigenvalues $\lambda_1, \ldots, \lambda_n$ are non-negative and (2.23) and (2.24) extend to all real $r, s \geq 0$, furthermore[2]

$$\operatorname{tr}^{1/s}(A^s) = (\lambda_1^s + \cdots + \lambda_n^s)^{1/s} \leq (\lambda_1^r + \cdots + \lambda_n^r)^{1/r} = \operatorname{tr}^{1/r}(A^r) \tag{2.25}$$

whenever $s \geq r > 0$. If $A$ is positive definite, then the eigenvalues are strictly positive and (2.23) and (2.24) are valid for all $r, s \in \mathbb{R}$.

(c) Let $A, B \in \mathbb{R}^{n\times n}$ be positive semi-definite. Then $\sqrt{A}$ is well-defined and symmetric and $\langle v, \sqrt{A}B\sqrt{A}v\rangle = \langle\sqrt{A}v, B\sqrt{A}v\rangle \geq 0$ for all $v \in \mathbb{R}^n$, because $B$ is positive semi-definite, hence $\sqrt{A}B\sqrt{A}$ is also positive semi-definite. Hence (2.21) and (2.24) applied to $\sqrt{A}B\sqrt{A}$ and $r = 1$ yield

$$\operatorname{tr}(AB) = \operatorname{tr}(\sqrt{A}B\sqrt{A}) \geq 0. \tag{2.26}$$

Applying (2.26), the Cauchy-Schwarz inequality (2.22) and (2.25) with $s = 2 \geq 1 = r$ shows that

$$0 \leq \operatorname{tr}(AB) \leq \sqrt{\operatorname{tr}(A^2)}\sqrt{\operatorname{tr}(B^2)} \leq \operatorname{tr}(A)\operatorname{tr}(B), \tag{2.27}$$

cf. [2, Theorem 1] and the references given there. We will use this result once for (2.38) below.

*Proof of Theorem 2.1.* The main idea of the proof is to construct a sequence $(f_k)_{k\in\mathbb{N}}$ of $C^2$-functions $f_k\colon \mathbb{R}^n \to [0,\infty)$ approximating the Euclidean norm $\mathbb{R}^n \ni z \mapsto |z|$, such that Itō's multi-dimensional formula can be applied to $f_k(X_t - Y_t)$, where $X$ and $Y$ are two solutions of the SDE (1.2) with $X_0 \overset{\text{a.s.}}{=} Y_0$. By then passing to the limit $k \to \infty$, the aim is to show that $\mathbb{E}[|X_t - Y_t|] = 0$ for all $t \geq 0$, which implies pathwise uniqueness.

The first step is to construct such approximations. Due to assumption (2.4), there exists a sequence

$$\gamma = a_0 > a_1 > a_2 > \cdots > a_k \searrow 0$$

such that

$$\int_{a_k}^{a_{k-1}} \frac{du}{\varrho^2(u)} = k, \quad k \in \mathbb{N}.$$

For every $k \in \mathbb{N}$, we can construct a continuous function $\phi_k\colon [0,\infty) \to [0,\infty)$ such that

$$\phi_k(u) \begin{cases} \leq \frac{2}{k\varrho^2(u)} & \text{for } u \in (a_k, a_{k-1}), \\ = 0 & \text{otherwise,} \end{cases} \tag{2.28}$$

and

$$\int_{a_k}^{a_{k-1}} \phi_k(u)\,du = 1,$$

because the upper bound (2.28) of $\phi_k$ integrates to 2 over $(a_k, a_{k-1})$. Next we define the auxiliary function $\varphi_k\colon [0,\infty) \to [0,\infty)$ by

$$\varphi_k(w) = \int_0^w \int_0^v \phi_k(u)\,du\,dv, \quad w \geq 0.$$

[2]Note that this well-known inequality is positive homogeneous in $(\lambda_1, \ldots, \lambda_n)$, hence it suffices to prove it for $\lambda_1^r + \cdots + \lambda_n^r = 1$, i.e. to show in this case that $\lambda_1^s + \cdots + \lambda_n^s \leq 1$. This is clear, because $\lambda_i \in [0,1]$, hence $\lambda_i^s \leq \lambda_i^r$ for each $i \in \{1, \ldots, n\}$.

Note that $\varphi_k$ is a twice continuously differentiable function with $\varphi_k(w) = 0$ for $w \in [0, a_k]$. Furthermore,

$$\varphi_k'(w) = \int_0^w \phi_k(u)\,du \begin{cases} = 0 & \text{for } w \in [0, a_k], \\ \le 1 & \text{for } w \in (a_k, a_{k-1}), \\ = 1 & \text{for } w \in [a_{k-1}, \infty). \end{cases} \tag{2.29}$$

Therefore, the sequence $(\varphi_k)_{k\in\mathbb{N}}$ is monotone increasing with $w - a_{k-1} \le \varphi_k(w) \le w$ for all $w \in [a_{k-1}, \infty)$. Finally, we can define the approximating sequence $(f_k)_{k\in\mathbb{N}}$ by

$$f_k(z) := \varphi_k(|z|), \quad k \in \mathbb{N},\, z \in \mathbb{R}^n.$$

It follows that each $f_k$ is a twice continuously differentiable function on $\mathbb{R}^n$ and that $f_k(z) \nearrow |z|$ uniformly in $z \in \mathbb{R}^n$ as $k \to \infty$.

Now, let $X$ and $Y$ be two solutions of (1.2) with $X_0 \overset{\text{a.s.}}{=} Y_0$, driven by the same $d$-dimensional Brownian motion, and define the difference process $Z$ by

$$Z_t := X_t - Y_t = \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s, \quad t \ge 0,$$

where we simplified the notation by defining the $\mathbb{R}^n$-valued stochastic process

$$b_s = b(s, X_s) - b(s, Y_s), \quad s \ge 0,$$

and the matrix-valued stochastic process

$$\sigma_s = \sigma(s, X_s) - \sigma(s, Y_s), \quad s \ge 0.$$

Define $\tau = \inf\{t \ge 0 : |Z_t| \ge \gamma\}$. Since $\{z \in \mathbb{R}^n : |z| \ge \gamma\}$ is closed and $Z$ has continuous paths, $\tau$ is a stopping time. By assumption (2.2),

$$|b_{s\wedge\tau}| \le \kappa(|Z_{s\wedge\tau}|), \quad s \ge 0. \tag{2.30}$$

We note that the definition of the Frobenius matrix norm and assumption (2.3) imply

$$\operatorname{tr}[\sigma_{s\wedge\tau}\sigma_{s\wedge\tau}^\top] = \|\sigma_{s\wedge\tau}\|_{\mathrm{F}}^2 \le \varrho^2(|X_{s\wedge\tau} - Y_{s\wedge\tau}|) = \varrho^2(|Z_{s\wedge\tau}|), \quad s \ge 0. \tag{2.31}$$

Fix $k \in \mathbb{N}$. Applying Itō's multi-dimensional formula to $f_k(Z_t)$, we obtain up to indistinguishability,

$$f_k(Z_t) = I_k(t) + J_k(t), \quad t \ge 0, \tag{2.32}$$

with

$$I_k(t) := \int_0^t \nabla f_k(Z_s)\,\sigma_s\,dW_s, \quad t \ge 0, \tag{2.33}$$

and

$$J_k(t) := \int_0^t \Bigl(\nabla f_k(Z_s)\,b_s + \frac{1}{2}\operatorname{tr}[H_k(Z_s)\sigma_s\sigma_s^\top]\Bigr)\,ds, \quad t \ge 0, \tag{2.34}$$

where $\nabla f_k(z)$ and $H_k(z)$ denote the gradient vector and the Hessian matrix of $f_k$ at $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$, respectively. We will now fix $t \ge 0$ and define suitable stopping times, so that we can treat the expectation of these two terms.

The stochastic process $I_k$ is a local martingale starting at zero, hence there exists an increasing sequence $(T_{k,l})_{l\in\mathbb{N}}$ of stopping times with $T_{k,l} \to \infty$ as $l \to \infty$ such that, for every $l \in \mathbb{N}$, the process $M_{k,l}(s) := I_k(s \wedge T_{k,l})$ with $s \ge 0$ is a uniformly integrable martingale. By Doob's optional stopping theorem [7, Ch. II, §3],

$$\mathbb{E}[I_k(t \wedge \tau \wedge T_{k,l})] = 0, \quad l \in \mathbb{N}. \tag{2.35}$$

Note that $\nabla f_k(z) = 0$ for $|z| \le a_k$ and $\nabla f_k(z) = \varphi_k'(|z|)z/|z|$ for $|z| \ge a_k$, hence $|\nabla f_k(z)| \le 1$ for all $z \in \mathbb{R}^n$ by (2.29). Therefore, by the Cauchy–Schwarz inequality and estimate (2.30),

$$|\nabla f_k(Z_{s\wedge\tau})\, b_{s\wedge\tau}| \le |\nabla f_k(Z_{s\wedge\tau})|\, |b_{s\wedge\tau}| \le \kappa(|Z_{s\wedge\tau}|), \quad s \ge 0. \tag{2.36}$$

The diagonal components of the Hessian matrix are given by

$$(H_k(z))_{i,i} = \frac{\partial^2 f_k(z)}{\partial z_i^2} = \phi_k(|z|)\frac{z_i^2}{|z|^2} + \varphi_k'(|z|)\frac{|z|^2 - z_i^2}{|z|^3}, \quad i \in \{1, \dots, n\},$$

(remember that $\varphi_k'$ and $\phi_k$ are zero in a neighborhood of the origin). Since $\varphi_k'$ is uniformly bounded by one, see (2.29), the above equation implies that

$$\operatorname{tr}[H_k(z)] = \phi_k(|z|) + \varphi_k'(|z|)\frac{n-1}{|z|} \le \phi_k(|z|) + \frac{n-1}{|z|}\mathbb{I}_{\{z\ne 0\}}, \quad z \in \mathbb{R}^n. \tag{2.37}$$

Note that $\varphi_k$ is convex on $[0,\infty)$ by construction, also the Euclidean norm is convex on $\mathbb{R}^n$, hence $f_k$ is convex. Therefore, the Hessian $H_k$ of $f_k$ is positive semi-definite everywhere. Since also $\sigma_{s\wedge\tau}\sigma_{s\wedge\tau}^\top$ is positive semi-definite, it follows from (2.27) that

$$0 \le \operatorname{tr}[H_k(Z_{s\wedge\tau})\sigma_{s\wedge\tau}\sigma_{s\wedge\tau}^\top] \le \operatorname{tr}[H_k(Z_{s\wedge\tau})]\operatorname{tr}[\sigma_{s\wedge\tau}\sigma_{s\wedge\tau}^\top], \quad s \ge 0. \tag{2.38}$$

Combining this inequality with (2.31) and (2.37) in the first step and using (2.28) in the second one implies that

$$\begin{aligned} 0 \le \operatorname{tr}[H_k(Z_{s\wedge\tau})\sigma_{s\wedge\tau}\sigma_{s\wedge\tau}^\top] &\le \phi_k(|Z_{s\wedge\tau}|)\varrho^2(|Z_{s\wedge\tau}|) + \frac{n-1}{|Z_{s\wedge\tau}|}\varrho^2(|Z_{s\wedge\tau}|)\mathbb{I}_{\{Z_{s\wedge\tau}\ne 0\}} \\ &\le \frac{2}{k} + \frac{n-1}{|Z_{s\wedge\tau}|}\varrho^2(|Z_{s\wedge\tau}|)\mathbb{I}_{\{Z_{s\wedge\tau}\ne 0\}}, \quad s \ge 0. \end{aligned} \tag{2.39}$$

Inserting the estimates (2.36) and (2.39) into (2.34) and using the upper bound (2.5) given by $G$, it follows that

$$|J_k(t\wedge\tau)| \le \frac{t}{k} + \int_0^{t\wedge\tau} G(|Z_{s\wedge\tau}|)\, ds, \quad t \ge 0.$$

It follows from (2.32) that, for all $l \in \mathbb{N}$ and $t \ge 0$,

$$\mathbb{E}[f_k(Z_{t\wedge\tau\wedge T_{k,l}})] = \mathbb{E}[I_k(t\wedge\tau\wedge T_{k,l})] + \mathbb{E}[J_k(t\wedge\tau\wedge T_{k,l})].$$

The first expectation on the right-hand side vanishes due to (2.35). Noting that $G$ is non-negative, it follows that

$$\mathbb{E}[f_k(Z_{t\wedge\tau\wedge T_{k,l}})] \le \frac{t}{k} + \int_0^t \mathbb{E}[G(|Z_{s\wedge\tau}|)]\, ds, \quad l \in \mathbb{N}, t \ge 0.$$

Since by assumption $G$ is concave on $[0,\gamma]$, Jensen's inequality implies that

$$\mathbb{E}[G(|Z_{s\wedge\tau}|)] \le G(\mathbb{E}[|Z_{s\wedge\tau}|]), \quad s \ge 0. \tag{2.40}$$

Letting $l \to \infty$, using Fatou's lemma, it follows that

$$\mathbb{E}[f_k(Z_{t\wedge\tau})] \le \frac{t}{k} + \int_0^t G(\mathbb{E}[|Z_{s\wedge\tau}|])\, ds.$$

Letting $k \to \infty$ and using monotone converge theorem, we obtain for the difference process $Z$ the estimate

$$\mathbb{E}[|Z_{t\wedge\tau}|] \le \int_0^t G(\mathbb{E}[|Z_{s\wedge\tau}|])\, ds, \quad t \ge 0. \tag{2.41}$$

The stopping at $\tau$ makes sure that $[0,\infty) \ni s \mapsto \mathbb{E}[|Z_{s\wedge\tau}|]$ is $[0,\gamma]$-valued and continuous (apply the dominated convergence theorem). Due to (2.6) and Bihari's inequality (see Theorem 3.1(b) below with $\beta \equiv 1$, $u(s) = \mathbb{E}[|Z_{s\wedge\tau}|]$ and $w(x) = G(x \wedge \gamma)$ for all $x \geq 0$), estimate (2.41) implies that $\mathbb{E}[|Z_{t\wedge\tau}|] = 0$. Since $Z_{t\wedge\tau} = Z_t 1_{\{\tau>t\}} + Z_\tau 1_{\{\tau\leq t\}}$ and $|Z_\tau| = \gamma > 0$ on $\{\tau < \infty\}$, it follows that $\mathbb{P}[\tau \leq t] = 0$, hence $\mathbb{E}[|Z_t|] = 0$, therefore $X_t \overset{\text{a.s.}}{=} Y_t$. Since this holds for all rational $t \geq 0$ and since the processes $X$ and $Y$ have continuous paths, they are indistinguishable. □

## 3. Bihari's inequality

Bihari's inequality [1, 6], proved by Hungarian mathematician Imre Bihari (1915–1998), is a nonlinear generalization of the Grönwall–Bellman inequality[3]. It is an important tool to obtain various estimates in the theory of ordinary and stochastic differential equations.

**Theorem 3.1.** *Let $I$ denote an interval of the real line of the form $[a,\infty)$, $[a,b]$ or $[a,b)$ with $a < b$. Let $\beta$, $u$: $I \to [0,\infty)$ and $w$: $[0,\infty) \to [0,\infty)$ be three functions, where $u$ and $w$ are continuous on $I$, $\beta$ is continuous on the interior $I^\circ$ of $I$ with $\int_a^t \beta(s)\,ds < \infty$ for all $t \in I$, and $w$ is non-decreasing and strictly positive on $(0,\infty)$.*

(a) *If, for some $\alpha > 0$, the function $u$ satisfies the inequality*

$$u(t) \leq \alpha + \int_a^t \beta(s)\, w(u(s))\, ds, \qquad t \in I, \tag{3.2}$$

*then*

$$u(t) \leq F^{-1}\Bigl(\int_a^t \beta(s)\, ds\Bigr), \qquad t \in [a,T), \tag{3.3}$$

*where $F^{-1}$ is the inverse function of*

$$F(x) := \int_\alpha^x \frac{dy}{w(y)}, \qquad x > 0,$$

*and*

$$T := \sup\Bigl\{ t \in I \Bigm| \int_a^t \beta(s)\, ds < \int_\alpha^\infty \frac{dy}{w(y)} \Bigr\}.$$

(b) *If the function $u$ satisfies the inequality* (3.2) *with $\alpha = 0$ and*

$$\int_0^x \frac{dy}{w(y)} = \infty \qquad \text{for all } x > 0, \tag{3.4}$$

*then $u(t) = 0$ for all $t \in I$.*

*Remark* 3.5. If $\int_\alpha^\infty \frac{dy}{w(y)} = \infty$, then (3.3) is valid on $[0,\infty)$. An example of such a function is $w(y) = y$ for $y \in [0,\infty)$.

*Remark* 3.6. The assumptions on the function $\beta$ allow for a singularity at the left end point $a$ of the interval $I$, for example $\beta(s) = (s-a)^{-\gamma}$ for $s > a$ with $\gamma \in (0,1)$. The integrability assumption for $\beta$ ensures that $T > a$ in (3.3).

[3] See also *Gronwall's inequality* at en.wikipedia.org/wiki/, version of December 8, 2012.

*Proof of Theorem 3.1.* (a) Denoting the right-hand side of (3.2) by

$$v(t) := \alpha + \int_a^t \beta(s)\, w(u(s))\, ds, \qquad t \in I,$$

we have $u \le v$ on $I$ by (3.2), which implies that $w(u(s)) \le w(v(s))$ for all $s \in I$ since $w$ is non-decreasing. Using $\alpha > 0$, the definitions of $F$ and $v$ as well as this inequality, it follows that

$$\frac{dF(v(s))}{ds} = \frac{v'(s)}{w(v(s))} = \frac{\beta(s)\, w(u(s))}{w(v(s))} \le \beta(s), \qquad s \in I^\circ.$$

Integrating this between $a$ and $t$ and using $F(v(a)) = F(\alpha) = 0$,

$$F(v(t)) = F(v(t)) - F(v(a)) \le \int_a^t \beta(s)\, ds, \qquad t \in I.$$

Since $F$ is strictly increasing,

$$v(t) \le F^{-1}\biggl(\int_a^t \beta(s)\, ds\biggr), \qquad t \in [a, T).$$

Since $u(t) \le v(t)$, the inequality (3.3) follows.

(b) Consider any $t \in I$ and $x > 0$. Due to (3.4) there exists $\alpha \in (0, x]$ such that

$$\int_\alpha^x \frac{dy}{w(y)} = \int_a^t \beta(s)\, ds.$$

Since $u$ also satisfies (3.2) with this $\alpha$, (3.3) implies that

$$u(t) \le F^{-1}\biggl(\int_a^t \beta(s)\, ds\biggr) = x.$$

Since $x > 0$ was arbitrary, $u(t) = 0$ follows. $\square$

## References

1. I. Bihari, *A generalization of a lemma of Bellman and its application to uniqueness problems of differential equations*, Acta Mathematica Hungarica **7** (1956), no. 1, 81–94, DOI: 10.1007/BF02022967. *[9]*
2. I. D. Coope, *On matrix trace inequalities and related topics for products of Hermitian matrices*, J. Math. Anal. Appl. **188** (1994), no. 3, 999–1001. *[6]*
3. R. Durrett, *Stochastic Calculus: A Practical Introduction*, CRC, 1996. *[1]*
4. N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North Holland, 1989. *[1]*
5. W. F. Osgood, *Beweis der Existenz einer Lösung der Differentialgleichung $\frac{dy}{dx} = f(x, y)$ ohne Hinzunahme der Cauchy-Lipschitz'schen Bedingung*, Monatshefte für Mathematik **9** (1898), 331–345, DOI: 10.1007/BF01707876. *[3]*
6. B. G. Pachpatte, *Inequalities for Differential and Integral Equations*, Academic Press, 1998. *[9]*
7. D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3. ed., vol. 293, Springer-Verlag, Berlin, 1999. *[1, 3, 7]*
8. G. Teschl, *Ordinary Differential Equations and Dynamical Systems*, Graduate Studies in Mathematics, vol. 140, American Mathematical Society, Providence, Rhode Island, 2012. *[2]*
9. T. Yamada and S. Watanabe, *On the uniqueness of solutions of stochastic differential equations*, J. Math. Kyoto Univ. **11** (1971), no. 1, 155–167. *[1]*
10. ———, *On the uniqueness of solutions of stochastic differential equations II*, J. Math. Kyoto Univ. **11** (1971), no. 3, 533–563. *[1, 3]*

(S. Altay) Department of Financial and Actuarial Mathematics, Vienna University of Technology, Wiedner Hauptstrasse 8-10, 1040 Vienna, Austria
*E-mail address*: altay@fam.tuwien.ac.at

(U. Schmock) Department of Financial and Actuarial Mathematics, Vienna University of Technology, Wiedner Hauptstrasse 8-10, 1040 Vienna, Austria
*E-mail address*: schmock@fam.tuwien.ac.at
*URL*: http://www.fam.tuwien.ac.at/~schmock/