STOCHASTIC EQUATION ON COMPACT GROUPS IN DISCRETE NEGATIVE TIME

JIRÔ AKAHORI, CHIHIRO UENISHI, AND KOUJI YANO

Dedicated to Professor Shinzo Watanabe on the occasion of his 70th birthday and to Professor Yoichiro Takahashi on the occasion of his 60th birthday

March 3, 2006

Abstract. In this paper a stochastic equation on compact groups in discrete negative time is studied. This is closely related to Tsirelson's stochastic differential equation, of which any solution is non-strong. How the group action reflects on the set of solutions is investigated. It is applied to generalize Yor's result and give a necessary and sufficient condition for existence of a strong solution and for uniqueness in law.

1. Introduction

In contrast with that of ordinary ones, the theory of stochastic differential equations has the distinguished notions of a *strong solution* and two uniqueness properties: *pathwise uniqueness* and *uniqueness in law*. Therefore we have the following four cases:

(1.1)

	unique in law	non-unique in law
$\exists$ strong solution	(**C0**)	(**C2**)
$\nexists$ strong solution	(**C1**)	(**C3**)

The celebrated theorem of T. Yamada and S. Watanabe ([14]) is stated as follows: *Pathwise uniqueness implies uniqueness in law and then any solution is strong.* Except for the trivial cases where there is no solution, we may say that pathwise uniqueness implies that the case (**C0**) occurs. In most cases the converse is also true (See A. K. Zvonkin and N. V. Krylov [16]).

B. Tsirelson ([12]) has presented a remarkable example of a stochastic differential equation (1.7) stated below, which enjoys uniqueness in law

2000 *Mathematics Subject Classification.* Primary 60H10, Secondary 60B15.

Key words and phrases. Stochastic differential equation; strong solution; uniqueness in law.

This research was supported by Open Research Center Project for Private Universities: matching fund subsidy from MEXT, 2004-2008.

property but has no strong solution: In short, the case (**C1**) occurs. His equation has deeply been investigated from various viewpoints by many researchers: See, for example, [1], [4] and [6].

To prove the non-existence of strong solutions, Tsirelson introduced a stochastic equation on the torus $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ indexed by discrete negative time:

$$\eta_k = \eta_{k-1} + \xi_k, \quad k \in -\mathbf{N}, \tag{1.2}$$

where $\xi = (\xi_k, k \in -\mathbf{N})$ is a Gaussian driving noise and $\eta = (\eta_k, k \in -\mathbf{N})$ is an unknown process. We can find its simpler proof, mainly due to N. V. Krylov, in the literature: For example, see [7, pp.150–151], [9, pp.149–151] and [2, pp.195–197]. In these contexts, the authors discussed, instead of (1.2), a modified equation on the real line $\mathbf{R}$ as follows:

$$\eta_k = \alpha(\eta_{k-1}) + \xi_k, \quad k \in -\mathbf{N}. \tag{1.3}$$

Here $\alpha(x)$ denotes the fractional part of $x \in \mathbf{R}$.

M. Yor ([15]) has studied the stochastic equation (1.3) for an *arbitrarily* given noise law, which is not necessarily Gaussian. In this case we always have a non-strong solution, so that the case (**C0**) never occurs. He successfully obtained striking results to give a necessary and sufficient condition for the trichotomy (**C1**)-(**C3**) in terms of the noise law. One of his results may roughly be stated as follows. To any given noise law μ there corresponds a subgroup $\mathbf{Z}_\mu$ of $\mathbf{Z}$ such that the following holds.

Theorem 1.1 (Yor [15])**.** (i) *The case* (**C1**) *occurs iff* $\mathbf{Z}_\mu = \{0\}$.
(ii) *The case* (**C2**) *occurs iff* $\mathbf{Z}_\mu = \mathbf{Z}$. *If it is true, then, among the solutions, all extremal points are strong and the others non-strong.*
(iii) *The case* (**C3**) *occurs iff* $\{0\} \subsetneq \mathbf{Z}_\mu \subsetneq \mathbf{Z}$.

This result sounds somehow paradoxical; at least, they are mysterious: *Uniqueness in law implies that the solution is non-strong, and existence of strong solution(s) implies that it is non-unique but extremal.*

The purpose of the present paper is to generalize Yor's results. We study the following stochastic equation on a compact group G in discrete negative time:

$$\eta_k = \eta_{k-1} \cdot \xi_k, \quad k \in -\mathbf{N}, \tag{1.4}$$

which we call the *simple Tsirelson–Yor equation on* G. We shall introduce the notions of a solution, a strong solution and uniqueness in law following the theory of stochastic differential equations. Then we obtain the following result, which generalizes Theorems 2 of [15].

Theorem 1.2. *For any given noise law μ, there exists a solution $\boldsymbol{P}_\mu^*$ of the equation* (1.4) *such that all marginal distributions of $\boldsymbol{P}_\mu^*$ are uniform. If $G \neq \{e\}$, then the solution $\boldsymbol{P}_\mu^*$ is non-strong.*

This result says that, if $G \neq \{e\}$, then the case (**C0**) always fails. Theorem 1.2 will be restated in Theorem 3.2

We shall naturally define a group action of G on the set of solutions. Our key result is roughly stated as follows.

Theorem 1.3. *The action restricted on the set of extremal points is transitive.*

This implies that the set of extremal points is exhausted by the G-orbit of an arbitrary extremal point $\boldsymbol{P}^o$ and is homeomorphic to the homogeneous space $G/H_\mu(\boldsymbol{P}^o)$ where $H_\mu(\boldsymbol{P}^o)$ is the isotropic subgroup at the point $\boldsymbol{P}^o$. The precise statement of Theorem 1.3 will be given in Theorem 4.3.

Based on Theorem 1.3 and fully employing the representation theory of compact groups, we generalize Theorem 1.1 to give a necessary and sufficient condition for the trichotomy (**C1**)-(**C3**) in terms of two subgroups $H_\mu(\boldsymbol{P}^o)$ and H_μ^s. Here H_μ^s is a closed normal subgroup of G which we will define in (5.3).

Theorem 1.4. (i) *The case* (**C1**) *occurs iff $H_\mu(\boldsymbol{P}^o) = G$ for some (and hence any) extremal point $\boldsymbol{P}^o$.*
(ii) *The case* (**C2**) *occurs iff $H_\mu^s = \{0\}$. If it is true, then, among the solutions, all extremal points are strong and the others non-strong.*
(iii) *The case* (**C3**) *occurs iff $H_\mu(\boldsymbol{P}^o) \subsetneq G$ and $H_\mu^s \supsetneq \{0\}$.*

Theorem 1.4 will be proved in section 7.

The subgroups H_μ^s and $H_\mu(\boldsymbol{P}^o)$ for extremal points $\boldsymbol{P}^o$ are related as follows.

Theorem 1.5. *For any extremal point $\boldsymbol{P}^o$, the inclusion*

$$H_\mu(\boldsymbol{P}^o) \subset H_\mu^s \tag{1.5}$$

holds. If G is abelian, then the equality holds for any extremal point $\boldsymbol{P}^o$.

We will restate Theorem 1.5 as Theorem 5.1 including further information. In the case where $G = \mathbf{T} = \mathbf{R}/\mathbf{Z}$, note that the abelian group $\mathbf{Z}$ is the dual group of G in the sense of the *Pontryagin duality.* In this terminology, the subgroup $\mathbf{Z}_\mu$ of $\mathbf{Z}$ in Theoren 1.1 is exactly the dual group of H_μ^s. In the non-abelian cases, however, we have a typical example given in Example 6.3 where $H_\mu(\boldsymbol{P}^o)$ is strictly included in H_μ^s for any extremal point $\boldsymbol{P}^o$.

Now we recall Tsirelson's stochastic differential equation (SDE in short). Let $(t_k : k \in -\mathbf{N})$ be a decreasing sequence of the interval $(0, 1]$ such that $t_k \to 0$ as $k \to -\infty$. Define $A(t, X)$, which is called *Tsirelson's drift*, as

$$(1.6) \qquad A(t,X) = \sum_{k \in -\mathbf{N}} \left\{ \frac{X_{t_{k-1}} - X_{t_{k-2}}}{t_{k-1} - t_{k-2}} \right\} 1_{[t_{k-1}, t_k)}(t), \quad 0 < t < 1.$$

Then Tsirelson's SDE is given by

$$(1.7) \qquad X_t = B_t + \int_0^t A(s, X) ds, \quad 0 \le t \le 1$$

where (B_t) is a one-dimensional Brownian motion. If $((X_t), (B_t))$ is a solution of the equation (1.7), then the sequences (η_k) and (ξ_k) defined by

$$(1.8) \qquad \eta_k = \frac{X_{t_k} - X_{t_{k-1}}}{t_k - t_{k-1}}, \quad \xi_k = \frac{B_{t_k} - B_{t_{k-1}}}{t_k - t_{k-1}}, \quad k \in -\mathbf{N}$$

satisfy the equation (1.3). Conversely, we can reconstruct the process (X_t) from the processes (η_k) and (B_t). Hence we may say that all properties of the SDE (1.7) can be deduced from those of the equation (1.3).

To generalize the stochastic equation (1.3), we introduce a class of stochastic equations in discrete negative time, which we call *Tsirelson–Yor equations*. Let S be a Polish space and G a compact group. We introduce an operation of G on the state space S through a measurable map $\psi : G \times S \to S$ and consider a measurable map $\theta : S \to G$. Then our Tsirelson–Yor equation is of the form

$$(1.9) \qquad \eta_k = \psi\left(\theta(\eta_{k-1}), \xi_k\right), \quad k \in -\mathbf{N}.$$

The key to the generalization is the following *commutation* condition[1]:

$$(1.10) \qquad \theta\left(\psi(g, s)\right) = g \cdot \theta(s), \quad g \in G, \ s \in S.$$

This condition allows us to reduce the equation (**??**) to the equation

$$(1.11) \qquad \theta(\eta_k) = \theta(\eta_{k-1}) \cdot \theta(\xi_k), \quad k \in -\mathbf{N}.$$

If we write $\widehat{\eta}_k = \theta(\eta_k)$ and $\widehat{\xi}_k = \theta(\xi_k)$, then the equation (1.11) is exactly the simple Tsirelson–Yor equation on G.

We will prove in Proposition 8.6 that in order to study the Tsirelson–Yor equation (1.11) with given laws it is sufficient to investigate the simple Tsirelson–Yor equation on G.

[1]The condition was discovered in [13].

We will show in Example 8.4 that Yor's equation (1.3) is an example of our Tsirelson–Yor equation. Proposition 8.6 says that the equation (1.3) is essentially 'equivalent' to the equation

$$\alpha(\eta_k) = \alpha(\eta_{k-1}) + \alpha(\xi_k) \quad \text{modulo } 1, \quad k \in -\mathbf{N}, \tag{1.12}$$

which is nothing but the equation (1.2).

In addition, we will show in Example 8.5 that the following equation is also an example:

$$\eta_k = \operatorname{sgn}(\eta_{k-1}) \cdot \xi_k, \quad k \in -\mathbf{N} \tag{1.13}$$

provided that the noise law has no point mass at $x = 0$. Here $\operatorname{sgn}(x) = 1$ if $x \geq 0$ and $\operatorname{sgn}(x) = -1$ if $x < 0$. This equation has been dealt with in [8, Chapter IX, Exercise 3.18] and [6, pp. 87–88]. Proposition 8.6 says that the equation (1.13) is equivalent to the simple Tsirelson–Yor equation on $\mathbf{Z}/2\mathbf{Z}$:

$$\operatorname{sgn}(\eta_k) = \operatorname{sgn}(\eta_{k-1}) \cdot \operatorname{sgn}(\xi_k), \quad k \in -\mathbf{N}. \tag{1.14}$$

This equation will be dealt with in Example 6.1.

The present paper is organized as follows. In section 2, we give the precise definition of the simple Tsirelson–Yor equation on a compact group and introduce the notions of a solution, a strong solution and uniqueness in law. We also prepare some preliminary facts about the set of solutions. Section 3 is devoted to the proof of Theorem 1.2. In section 4, we give a precise statement and the proof of Theorem 1.3. In section 5, we introduce a subgroup H_μ^s which appears in Theorem 1.4 and restate Theorem 1.5 as Theorem 5.1. In section 6, we investigate the equation (1.14) in detail and consider another typical example of the simple Tsirelson–Yor equation. Section 7 is devoted to the proofs of Theorem 1.4 and the propositions which are given in section 5. The Tsirelson–Yor equations are defined and discussed in section 8.

Acknowledgments: The authors wish to express sincere thanks to Professors Marc Yor, Freddy Delbaen, Shinzo Watanabe, and Yoichiro Takahashi for stimulating discussions and valuable comments. They also thank Gen Mano and Hidehisa Alikawa, who kindly informed them of basic theorems of the representation theory. The second author expresses his hearty thanks to Professors Hiroki Aoki and Yoshiaki Kobayashi for their kind tutorial lectures.

2. Definitions and preliminary facts

Let G be a compact group. We consider the following stochastic equation on G in discrete negative time:

$$\eta_k = \eta_{k-1} \cdot \xi_k, \quad k \in -\mathbf{N}. \tag{1.4}$$

Definition 2.1. The stochastic equation (1.4) is called the *simple Tsirelson-Yor equation* on G, which will be abbreviated by "STYE".

For an arbitrary sequence $\mu = (\mu_k, k \in -\mathbf{N})$ of Borel probability measures μ_k on G, we consider the equation (1.4) with a general noise law μ.

Let us give the precise definition of a solution of (1.4). We denote by $\eta = (\eta_k, k \in -\mathbf{N})$ the coordinate mapping process: $\eta_k(\omega) = \omega(k)$ for $\omega \in G^{-\mathbf{N}}$ and $k \in -\mathbf{N}$. Set

$$\xi_k = (\eta_{k-1})^{-1} \cdot \eta_k, \quad k \in -\mathbf{N}. \tag{2.1}$$

Let $\mathcal{F}_k^\eta$ and $\mathcal{F}_k^\xi$ for $k \in -\mathbf{N}$ denote

$$\mathcal{F}_k^\eta := \sigma(\eta_k, \eta_{k-1}, \ldots), \quad k \in -\mathbf{N} \tag{2.2}$$

and

$$\mathcal{F}_k^\xi := \sigma(\xi_k, \xi_{k-1}, \ldots), \quad k \in -\mathbf{N} \tag{2.3}$$

respectively. It is obvious that

$$\mathcal{F}_k^\xi \subset \mathcal{F}_k^\eta, \quad k \in -\mathbf{N}. \tag{2.4}$$

Since the law on $G^{-\mathbf{N}}$ of the process η determines that of the noise ξ, the following definition is reasonable.

Definition 2.2. Let $\mu = (\mu_k, k \in -\mathbf{N})$ be a sequence of Borel probability measures μ_k on G. A *solution* of the STYE on G with the noise law μ is a probability measure $\boldsymbol{P}$ on $G^{-\mathbf{N}}$ such that the following two statements hold:
(i) ξ_k is independent of $\mathcal{F}_{k-1}^\eta$ under $\boldsymbol{P}$, for any $k \in -\mathbf{N}$.
(ii) ξ_k is distributed as μ_k under $\boldsymbol{P}$, for any $k \in -\mathbf{N}$.

The totality of solutions of the STYE with the noise law μ will be denoted by $\mathcal{P}_\mu$.

Remark 2.3. For a given μ, a probability measure $\boldsymbol{P}$ on $G^{-\mathbf{N}}$ belongs to $\mathcal{P}_\mu$ if and only if the following (inhomogeneous) Markov property holds:

$$\boldsymbol{E}\left[\phi(\eta_k) \mid \mathcal{F}_{k-1}\right] = \int_G \phi(\eta_{k-1} \cdot g)\mu_k(dg), \quad k \in -\mathbf{N} \tag{2.5}$$

for any bounded measurable function ϕ on G.

In order to give a precise meaning to the table (1.1), we need to introduce the notions of strong solution and uniqueness in law. We follow the usual terminology in the theory of stochastic differential equations.

Definition 2.4. Let μ be given. A solution $\boldsymbol{P} \in \mathcal{P}_\mu$ is called *strong* if

$$\mathcal{F}_k^\eta \subset \mathcal{F}_k^\xi \text{ up to } \boldsymbol{P}\text{-null sets}, \quad k \in -\mathbf{N}. \tag{2.6}$$

If $\boldsymbol{P} \in \mathcal{P}_\mu$ is not strong, then it is called *non-strong*[2].

Definition 2.5. Let μ be given. It is said that *uniqueness in law* holds if the set $\mathcal{P}_\mu$ consists of at least one element.

Let $\mu = (\mu_k, k \in -\mathbf{N})$ be given and consider the set $\mathcal{P}_\mu$ of solutions of the STYE on G with the noise law μ. Since the condition $\boldsymbol{P} \in \mathcal{P}_\mu$ is equivalent to the Markov property (2.5), the following is obvious.

Lemma 2.6. *The set $\mathcal{P}_\mu$ is a closed convex subset of the linear topological space of signed measures on $G^{-\mathbf{N}}$ equipped with the weak-star topology.*

We denote the totality of extremal points of $\mathcal{P}_\mu$ by $\mathrm{ex}(\mathcal{P}_\mu)$. Then Lemma 2.6 implies that $\mathrm{ex}(\mathcal{P}_\mu) \subset \mathcal{P}_\mu$ and any solution $\boldsymbol{Q} \in \mathcal{P}_\mu$ has an integral representation

$$\boldsymbol{Q}(\cdot) = \int_{\mathrm{ex}(\mathcal{P}_\mu)} \boldsymbol{P}(\cdot)\, \gamma(d\boldsymbol{P}) \tag{2.7}$$

for some Borel probability measure γ on $\mathrm{ex}(\mathcal{P}_\mu)$.

The following is also true.

Lemma 2.7. *For $\boldsymbol{P} \in \mathcal{P}_\mu$, the solution $\boldsymbol{P}$ is extremal if and only if $\mathcal{F}_{-\infty}^\eta$ is $\boldsymbol{P}$-trivial.*

We omit the proof, since the proof of [15, Theorem 1, 2)] still survives for the STYE's.

3. Existence of a non-strong solution

Since the group G is compact, we have the normalized Haar measure ν: That is, there exists a unique positive Borel measure ν on G such that $\nu(G) = 1$ and such that

$$\int_G \phi(gh)\nu(dh) = \int_G \phi(hg)\nu(dh) = \int_G \phi(h)\nu(dh) \tag{3.1}$$

for any $g \in G$ and any bounded measurable function ϕ on G.

[2] It is sometimes called *weak*.

Definition 3.1. A random variable U is said to be *uniformly distributed on G* if the law of U is equal to the normalized Haar measure ν.

Now we restate Theorem 1.2.

Theorem 3.2. *For an arbitrary noise law $\mu = (\mu_k)$, there exists a unique element $\boldsymbol{P}^*_\mu \in \mathcal{P}_\mu$ whose marginal distributions are uniform:*

$$\lambda_k = \nu, \quad k \in -\mathbf{N}. \tag{3.2}$$

Moreover, under $\boldsymbol{P}^*_\mu$,

$$\text{each } \eta_k \text{ is independent of the noise } \mathcal{F}^\xi = \sigma(\xi_k : k \in -\mathbf{N}). \tag{3.3}$$

In particular, the solution $\boldsymbol{P}^*_\mu$ is non-strong *unless G is trivial: i.e. $G = \{e\}$.*

Theorem 3.2 immediately implies the following.

Corollary 3.3. *Assume that G is not trivial. Then, for any noise law μ, the case* (**C0**) *in the table* (1.1) *never occurs.*

In the sequel, we always assume that G is not trivial. Now the problem is how to characterize the trichotomy (**C1**)-(**C3**).

The key fact to the proof of Theorem 3.2 is the one-to-one and onto correspondence of each solution to what we call an entrance law.

Let $\mu = (\mu_k, k \in -\mathbf{N})$ be given.

Definition 3.4. A family $\lambda = (\lambda_k, k \in -\mathbf{N})$ of probability laws λ_k on G is called an *entrance law* for the noise law μ if the following recurrence relation holds:

$$\lambda_k = \lambda_{k-1} * \mu_k, \quad k \in -\mathbf{N}. \tag{3.4}$$

Here $\mu * \lambda$ for two measures μ and λ on G stands for the convolution of μ and λ:

$$\int_G \phi(g)\mu * \lambda(dg) = \int_G \int_G \phi(gh)\mu(dg)\lambda(dh) \tag{3.5}$$

for any bounded measurable function ϕ on G.

For any $\boldsymbol{P} \in \mathcal{P}_\mu$, let $\lambda = (\lambda_k, k \in -\mathbf{N})$ denote the marginal distributions of η_k on G , i.e.,

$$\lambda_k(\cdot) = \boldsymbol{P}(\eta_k \in \cdot), \qquad k \in -\mathbf{N}. \tag{3.6}$$

Then the equation (1.4) implies that (3.4) holds. Conversely, the following holds.

Lemma 3.5. *Let μ be a given noise law. Let $\lambda = (\lambda_k, k \in -\mathbf{N})$ be an entrance law for the noise law μ: The family $(\lambda_k, k \in -\mathbf{N})$ satisfies the consistency condition* (3.4)*. Then there exists a unique element $\boldsymbol{P} \in \mathcal{P}_\mu$ such that* (3.6) *holds.*

This is obvious by Kolmogorov's extention theorem, so we omit the proof.

Proof of Theorem 3.2. By Lemma 3.5, we see that there exists a solution $\boldsymbol{P}^*_\mu$ with marginal distributions given by (3.2). Let $\boldsymbol{E}^*_\mu$ denote the expectation with respect to $\boldsymbol{P}^*_\mu$. Let ϕ be a bounded measurable function on $G \times G^{\{k-n+1,\dots,0\}}$. Noting that the independence of η_{k-n} and $\sigma(\xi_{k-n+1}, \dots, \xi_0)$, we have

$$\begin{aligned} &\boldsymbol{E}^*_\mu \left[\phi\left(\eta_k; \xi_{k-n+1}, \dots, \xi_0\right)\right] && (3.7)\\ &= \boldsymbol{E}^*_\mu \left[\phi\left(\eta_{k-n} \cdot \xi_{k-n+1} \cdots \xi_k; \xi_{k-n+1}, \dots, \xi_0\right)\right] && (3.8)\\ &= \boldsymbol{E}^*_\mu \left[\phi\left(\eta_{k-n}; \xi_{k-n+1}, \dots, \xi_0\right)\right]. && (3.9)\end{aligned}$$

This implies (3.3). □

4. Homogeneous space structure of the set of extremal points

Let $\mu = (\mu_k, k \in -\mathbf{N})$ be given. For each $g \in G$, we define a continuous map T_g on $\mathcal{P}_\mu$ by

$$T_g(\boldsymbol{P})(\cdot) = \boldsymbol{P}\left(g \cdot \eta \in \cdot\right), \quad \boldsymbol{P} \in \mathcal{P}_\mu \tag{4.1}$$

where we write

$$g \cdot \eta = \left(g \cdot \eta_k, k \in -\mathbf{N}\right), \quad g \in G,\ \eta \in G^{-\mathbf{N}}. \tag{4.2}$$

Then the following is obvious.

Lemma 4.1. *The family of maps $(T_g : g \in G)$ defines a group action of G on $\mathcal{P}_\mu$ with $\mathrm{ex}(\mathcal{P}_\mu)$ an invariant subset, i.e.,*

$$T_g(\mathcal{P}_\mu) \subset \mathcal{P}_\mu \quad \text{and} \quad T_g(\mathrm{ex}(\mathcal{P}_\mu)) \subset \mathrm{ex}(\mathcal{P}_\mu) \quad \text{for } g \in G \tag{4.3}$$

and

$$T_g T_{h^{-1}} = T_{gh^{-1}} \quad \text{for } g, h \in G. \tag{4.4}$$

The following proposition shows that the point $\boldsymbol{P}^*_\mu$, which is defined in Theorem 3.2, can be considered to be the *center* of the set of solutions $\mathcal{P}_\mu$.

Proposition 4.2. *Let μ be given. Then, for any $\boldsymbol{P} \in \mathcal{P}_\mu$, it holds that*

$$\boldsymbol{P}^*_\mu = \int_G T_g(\boldsymbol{P}) \nu(dg). \tag{4.5}$$

Proof of Proposition 4.2. For any bounded measurable function ϕ on G, we have

$$\int_G \phi(h) \int_G T_g(\boldsymbol{P})(\eta_k \in dh)\nu(dg) \tag{4.6}$$

$$= \int_G \boldsymbol{E}\left[\phi(g \cdot \eta_k)\right] \nu(dg) = \int_G \phi(g)\nu(dg) \tag{4.7}$$

for any $k \in -\mathbf{N}$. This implies that all marginal distributions of the RHS of (4.5) are uniform on G. □

Now we are in a position to give the precise statement of Theorem 1.3.

Theorem 4.3. *Let μ be given. Then the action of $(T_g : g \in G)$ restricted on $\mathrm{ex}(\mathcal{P}_\mu)$ is transitive: That is, if $\boldsymbol{P}^1$ and $\boldsymbol{P}^2$ are two solutions in $\mathrm{ex}(\mathcal{P}_\mu)$, then there exists an element $g \in G$ such that $\boldsymbol{P}^1 = T_g(\boldsymbol{P}^2)$.*

Let an extremal point $\boldsymbol{P}^o \in \mathrm{ex}(\mathcal{P}_\mu)$ be fixed. We denote by $H_\mu(\boldsymbol{P}^o)$ the isotropic subgroup at $\boldsymbol{P}^o$:

$$H_\mu(\boldsymbol{P}^o) = \{g \in G : T_g(\boldsymbol{P}^o) = \boldsymbol{P}^o\}. \tag{4.8}$$

It is easy to see that $H_\mu(\boldsymbol{P}^o)$ is a closed subgroup of G, and hence the quotient set $G/H_\mu(\boldsymbol{P}^o)$ is a compact Hausdorff set. Thus Theorem 4.3 implies the following corollary, which reveals the homogeneous space structure of the set $\mathrm{ex}(\mathcal{P}_\mu)$.

Corollary 4.4. *The action of $(T_g : g \in G)$ restricted on $\mathrm{ex}(\mathcal{P}_\mu)$ induces a homeomorphism:*

$$G/H_\mu(\boldsymbol{P}^o) \xrightarrow{\sim} \mathrm{ex}(\mathcal{P}_\mu). \tag{4.9}$$

The key to the proof of Theorem 4.3 is the following lemma, which we follow the proof of Yamada–Watanabe's theorem [14, Proposition 1]. See also [2, pp. 163–166].

We consider a product space $G^{-\mathbf{N}} \times G^{-\mathbf{N}} \times G^{-\mathbf{N}}$ with its coordinate written as (η^1, η^2, ξ). Define

$$\mathcal{F}_k^{\eta^1,\eta^2} = \sigma\left(\eta_j^1, \eta_j^2 : j \le k\right), \quad k \in -\mathbf{N}. \tag{4.10}$$

Lemma 4.5. *Let $\boldsymbol{P}^1$ and $\boldsymbol{P}^2$ be two solutions in $\mathcal{P}_\mu$. Then there exists a probability measure $\boldsymbol{Q}$ on $G^{-\mathbf{N}} \times G^{-\mathbf{N}} \times G^{-\mathbf{N}}$ such that the following statements hold:*
(i) For $i = 1, 2$, the law on $G^{-\mathbf{N}} \times G^{-\mathbf{N}}$ of (η^i, ξ) under $\boldsymbol{Q}$ coincides with that of (η, ξ) under $\boldsymbol{P}^i$ where ξ is defined in (2.1).
(ii) For any $k \in -\mathbf{N}$,

$$\xi_k \text{ is independent of } \mathcal{F}_{k-1}^{\eta^1,\eta^2} \text{ under } \boldsymbol{Q}. \tag{4.11}$$

Let $\boldsymbol{P}^1_\xi(\cdot)$ and $\boldsymbol{P}^2_\xi(\cdot)$ denote the regular conditional probability given $\mathcal{F}^\xi = \sigma(\xi_k : k \in -\mathbf{N})$ such that

$$\boldsymbol{P}^i(\eta \in A, \xi \in B) = \int_B \boldsymbol{P}^i_\xi(A)\mu(d\xi) \tag{4.12}$$

for arbitrary measurable sets A and B of $G^{-\mathbf{N}}$, $i = 1, 2$. Then the desired probability measure $\boldsymbol{Q}$ is obtained as

$$\boldsymbol{Q}(d\eta^1\, d\eta^2\, d\xi) = \boldsymbol{P}^1_\xi(d\eta^1)\, \boldsymbol{P}^2_\xi(d\eta^2)\, \mu(d\xi). \tag{4.13}$$

We can prove the claims (i) and (ii) in the same way as in the proof of Yamada–Watanabe's theorem, so we omit the proofs of Lemma 4.5.

Proof of Theorem 4.3. Let $\boldsymbol{Q}$ be the probability measure given in Lemma 4.5. As far as the end of this paragraph, we omit to write "$\boldsymbol{Q}$-a.s." By Lemma 4.5 (i), we have

$$\eta^i_k = \eta^i_{k-1} \cdot \xi_k, \quad k \in -\mathbf{N},\ i = 1, 2. \tag{4.14}$$

Let $k \in -\mathbf{N}$ and $n \in \mathbf{N}$ be arbitrary numbers. Then

$$\eta^i_k = \eta^i_{k-n} \cdot \xi_{k-n+1} \cdots \xi_k, \quad i = 1, 2. \tag{4.15}$$

Thus we have

$$\left(\eta^1_k\right) \cdot \left(\eta^2_k\right)^{-1} = \left(\eta^1_{k-n}\right) \cdot \left(\eta^2_{k-n}\right)^{-1}, \quad k \in -\mathbf{N},\ n \in \mathbf{N}. \tag{4.16}$$

Since the LHS is irrelevant to $n \in \mathbf{N}$, there exists a random variable ε which is $\mathcal{F}^{\eta^1,\eta^2}_{-\infty}$-measurable such that

$$\varepsilon = \left(\eta^1_k\right) \cdot \left(\eta^2_k\right)^{-1}, \quad k \in -\mathbf{N}. \tag{4.17}$$

If we denote by $\boldsymbol{Q}\,(\cdot \mid \varepsilon = g)$ the regular conditional probability given $\varepsilon = g$, then we have the following disintegration:

$$\boldsymbol{Q}(\cdot) = \int_G \boldsymbol{Q}\,(\cdot \mid \varepsilon = g)\, \boldsymbol{Q}(\varepsilon \in dg). \tag{4.18}$$

Hence we obtain an integral expression of $\boldsymbol{P}^1$:

$$\boldsymbol{P}^1(\cdot) = \int_G \boldsymbol{Q}\left(\eta^1 \in \cdot \mid \varepsilon = g\right) \boldsymbol{Q}(\varepsilon \in dg). \tag{4.19}$$

By definition, we see that the law

$$\boldsymbol{Q}\left(\eta^1 \in \cdot \mid \varepsilon = g\right) \tag{4.20}$$

belongs to $\mathcal{P}_\mu$ for $\boldsymbol{Q}(\varepsilon \in dg)$-almost every $g \in G$. By the assumption that $\boldsymbol{P}^1$ is extremal, we obtain

$$\boldsymbol{P}^1(\cdot) = \boldsymbol{Q}\left(\eta^1 \in \cdot \mid \varepsilon = g\right) \tag{4.21}$$

for $\boldsymbol{Q}(\varepsilon \in dg)$-almost every $g \in G$. We obtain the similar identity for $\boldsymbol{P}^2$. Note that (4.17) implies

$$\boldsymbol{Q}\left(\eta^1 = g \cdot \eta^2 \mid \varepsilon = g\right) = 1 \quad \text{for any } g \in G. \tag{4.22}$$

Therefore we conclude that

$$\boldsymbol{P}^1 = T_g\left(\boldsymbol{P}^2\right) \quad \text{for } \boldsymbol{Q}(\varepsilon \in dg)\text{-a.e. } g \in G. \tag{4.23}$$

This completes the proof. □

Remark 4.6. In the above proof of Theorem 4.3, take $g_0 \in G$ such that $\boldsymbol{P}^1 = T_{g_0}\left(\boldsymbol{P}^2\right)$. Then it is obvious that the measure $\boldsymbol{Q}(g_0^{-1}\varepsilon \in dg)$ is the normalized Haar measure on $H_\mu(\boldsymbol{P}^2)$. This fact leads to the following paradox: *The tail σ-field $\mathcal{F}_{-\infty}^{\eta^1,\eta^2}$ is non-trivial under $\boldsymbol{Q}$, whereas both $\mathcal{F}_{-\infty}^{\eta^1}$ and $\mathcal{F}_{-\infty}^{\eta^2}$ are trivial.*

Remark 4.7. Consider the STYE on the direct product group $G \times G$

$$(\eta_k^1, \eta_k^2) = (\eta_{k-1}^1, \eta_{k-1}^2) \cdot (\xi_k^1, \xi_k^2), \quad k \in -\mathbf{N} \tag{4.24}$$

with the noise law $\widetilde{\mu} = (\widetilde{\mu}_k : k \in -\mathbf{N})$ given by

$$\int_{G\times G} \phi(g,h)\widetilde{\mu}_k(dg \times dh) = \int_G \phi(g,g)\mu_k(dg) \tag{4.25}$$

for any bounded measurable function ϕ on $G \times G$ and for $k \in -\mathbf{N}$. Denote the set of solutions by $\widetilde{\mathcal{P}}_{\widetilde{\mu}}$. We denote the marginal laws of (η^1, η^2) under the measure $\boldsymbol{Q}$ by $\widetilde{\boldsymbol{P}}$. Then $\widetilde{\boldsymbol{P}}$ and the regular conditional probabilities $\widetilde{\boldsymbol{P}}(\cdot \mid \varepsilon = g)$ belong to $\widetilde{\mathcal{P}}_{\widetilde{\mu}}$. Note that (4.18) implies that

$$\widetilde{\boldsymbol{P}}(\cdot) = \int_G \widetilde{\boldsymbol{P}}(\cdot \mid \varepsilon = g)\boldsymbol{Q}(\varepsilon \in dg). \tag{4.26}$$

In this integral expression of the solution $\widetilde{\boldsymbol{P}}$, the integrand $\widetilde{\boldsymbol{P}}(\cdot \mid \varepsilon = g)$ belongs to $\mathrm{ex}(\widetilde{\mathcal{P}}_{\widetilde{\mu}})$ for $\boldsymbol{Q}(\varepsilon \in dg)$-almost every $g \in G$. In fact, by (4.21) and (4.22), we see that $\mathcal{F}_{-\infty}^{\eta^1,\eta^2}$ is $\widetilde{\boldsymbol{P}}(\cdot \mid \varepsilon = g)$-trivial for $\boldsymbol{Q}(\varepsilon \in dg)$-almost every $g \in G$.

5. The subgroup H_μ^s

To begin with, we recall the well-known *Peter–Weyl theorem* for compact groups (see, e.g., [10, Chapter 1] and [3, Corollary 13]).

Let G be a compact group and let ν denote the normalized Haar measure on G. Let $\mathcal{G}$ denote the totality of irreducible unitary representations ρ of G on a finite-dimensional linear space V^ρ. Then the following holds: *The family*

$$(\rho_{i,j} : 1 \le i, j \le \dim \rho,\ \rho \in \mathcal{G}) \tag{5.1}$$

forms a total family in the space of continuous functions on G. Here $(\rho_{i,j})$ denotes the matrix element of a representation $\rho \in \mathcal{G}$.

In what follows we consider the STYE (1.4) on a compact group G with a fixed noise law μ.

To characterize the trichotomy, we introduce a subset H^s_μ of G as follows. For an extremal point $\boldsymbol{P}^o \in \mathrm{ex}(\mathcal{P}_\mu)$ we define

$$\mathcal{H}^s_\mu(\boldsymbol{P}^o) = \Big\{\rho \in \mathcal{G} :\ \rho(\eta_k) \text{ is } \mathcal{F}^\xi_k\text{-m'ble } \boldsymbol{P}^o\text{-a.s. for } k \in -\mathbf{N}\Big\}. \tag{5.2}$$

Here the word "m'ble" is abbreviated from "measurable". It is clear from Theorem 4.3 that the set $\mathcal{H}^s_\mu(\boldsymbol{P}^o)$ is independent of the choice of $\boldsymbol{P}^o \in \mathrm{ex}(\mathcal{P}_\mu)$. So we simply write $\mathcal{H}^s_\mu(\boldsymbol{P}^o)$ as $\mathcal{H}^s_\mu$, and define

$$H^s_\mu = \{g \in G :\ \rho(g) = \mathrm{id} \text{ for every } \rho \in \mathcal{H}^s_\mu\}. \tag{5.3}$$

Note that we need to know at least one extremal point $\boldsymbol{P}^o$ in order to compute $H_\mu(\boldsymbol{P}^o)$ and H^s_μ. Let us introduce two subsets H^1_μ and H^2_μ of G which can directly be computed from the noise law μ as follows. For $\rho \in \mathcal{G}$, we set[3]

$$R_k = \int_G \rho(g)\mu_k(dg), \quad k \in -\mathbf{N}. \tag{5.4}$$

Then the following two limits exist for any $k \in -\mathbf{N}$: The first one is

$$r^1_k[\rho] = \lim_{n\to\infty} \|R_{k-n}R_{k-n+1}\cdots R_k\| \tag{5.5}$$

where $\|\cdot\|$ denotes the operator norm of linear operators on the representation space V^ρ. The second one is

$$r^2_k[\rho] = \lim_{n\to\infty} \left|\det\left(R_{k-n}R_{k-n+1}\cdots R_k\right)\right|. \tag{5.6}$$

The convergence of the first limit is obvious by $\|R_j\| \le 1$ for any $j \in -\mathbf{N}$. That of the second is ensured by $|\det R_j| \le 1$ for any $j \in -\mathbf{N}$, which will be assured by Lemma 7.4. Now we set

$$\mathcal{H}^i_\mu := \{\rho \in \mathcal{G} :\ r^i_k[\rho] > 0 \text{ for some } k \in -\mathbf{N}\}, \quad i = 1, 2 \tag{5.7}$$

and define

$$H^i_\mu := \{g \in G :\ \rho(g) = \mathrm{id} \text{ for every } \rho \in \mathcal{H}^i_\mu\}, \quad i = 1, 2, \tag{5.8}$$

where the symbol 'id' stands for the identity on the representation space V^ρ.

The following hierarchy is fundamental to our analysis.

---

[3]Here the integral in RHS of (5.4) is interpreted as the componentwise integral in a fixed matrix representation of ρ.

Theorem 5.1. (i) *The three subclasses $\mathcal{H}_\mu^1$, $\mathcal{H}_\mu^2$ and $\mathcal{H}_\mu^s$ satisfy the following inclusions:*

$$\mathcal{H}_\mu^1 \supset \mathcal{H}_\mu^s \supset \mathcal{H}_\mu^2. \tag{5.9}$$

(ii) *The three subsets H_μ^1, H_μ^2 and H_μ^s are closed normal subgroups of G such that*

$$H_\mu^1 \subset H_\mu(\boldsymbol{P}^o) \subset H_\mu^s \subset H_\mu^2 \tag{5.10}$$

for any $\boldsymbol{P}^o \in \mathrm{ex}(\mathcal{P}_\mu)$. If G is abelian, then the equalities hold:

$$H_\mu^1 = H_\mu(\boldsymbol{P}^o) = H_\mu^s = H_\mu^2. \tag{5.11}$$

We remark that this result includes the whole statement of Theorem 1.5. The proof of Theorem 5.1 will be given in section 7.

Remark 5.2. The isotropic subgroup $H_\mu(\boldsymbol{P})$ at another extremal point $\boldsymbol{P} = T_g(\boldsymbol{P}^o) \in \mathrm{ex}(\mathcal{P}_\mu)$ is related to $H_\mu(\boldsymbol{P}^o)$ by

$$H_\mu(\boldsymbol{P}) = gH_\mu(\boldsymbol{P}^o)g^{-1}. \tag{5.12}$$

Hence the isotropic subgroup $H_\mu(\boldsymbol{P}^o)$ is not necessarily normal, while the subgroup H_μ^s is always normal.

6. Examples

Example 6.1. Consider the STYE on the group $\mathbf{Z}/2\mathbf{Z} \simeq \{1, -1\}$. Since the group $\mathbf{Z}/2\mathbf{Z}$ is abelian, we have the equalities (5.11). Note that the class $\mathcal{G}$ consists of only one element ρ such that

$$\rho(1) = 1, \quad \rho(-1) = -1. \tag{6.1}$$

For a noise law $\mu = (\mu_k : k \in -\mathbf{N})$, we set $p_k = \mu_k(\{1\})$. Now set

$$r_k = \lim_{n\to\infty} \prod_{j=k-n}^{k} |2p_j - 1|, \quad k \in -\mathbf{N}. \tag{6.2}$$

Then Theorem 1.4 leads to the following.

Proposition 6.2. *The case* (**C1**) *or* (**C2**) *occurs according to whether the infinite product r_k vanishes for any $k \in -\mathbf{N}$ or not.*

This is obvious, so we omit the proof.

We give a typical example of the STYE on a non-abelian group where $H_\mu(\boldsymbol{P}^o)$ is non-normal and hence strictly included in H_μ^s.

Example 6.3. Consider the symmetric group of degree 3:

$$\mathfrak{S}_3 = \{e, (12), (23), (13), (123), (132)\}. \tag{6.3}$$

Set

$$H^o = \{e, (12)\}\,, \tag{6.4}$$

$$H^1 = (13)H^o = \{(13), (23)\}\,, \tag{6.5}$$

$$H^2 = (123)H^o = \{(123), (132)\}\,. \tag{6.6}$$

Then H^o is a non-normal subgroup of $\mathfrak{S}_3$ such that

$$\mathfrak{S}_3/H^o = \left\{H^o, H^1, H^2\right\}. \tag{6.7}$$

Let $\mu = (\mu_k : k \in -\mathbf{N})$ be the sequence of the uniform laws on H^o: $\mu_k = \nu^o$ for any $k \in -\mathbf{N}$ where

$$\nu^o(\{e\}) = \nu^o(\{(12)\}) = 1/2. \tag{6.8}$$

Proposition 6.4. *Consider the STYE on $\mathfrak{S}_3$ with the noise law μ given above. Then there exists a solution $\boldsymbol{P}^o \in \mathcal{P}_\mu$ such that the following hold:*
(i) *Each η_k under $\boldsymbol{P}^o$ is uniformly distributed on H^o.*
(ii) *The extremal points* $\mathrm{ex}(\mathcal{P}_\mu) = \left\{\boldsymbol{P}^o, \boldsymbol{P}^1, \boldsymbol{P}^2\right\}$, *where*

$$\boldsymbol{P}^1 := T_{(13)}(\boldsymbol{P}^o), \quad \boldsymbol{P}^2 := T_{(123)}(\boldsymbol{P}^o). \tag{6.9}$$

(iii) *The isotropic subgroup $H_\mu(\boldsymbol{P}^o) = H^o$. Hence*

$$\{e\} = H^1_\mu \subsetneq H_\mu(\boldsymbol{P}^o) = H^o \subsetneq H^s_\mu = H^2_\mu = \mathfrak{S}_3. \tag{6.10}$$

Proof. Note that the family $\mu = (\mu_k : k \in -\mathbf{N})$ itself forms an entrance law: $\nu^o = \nu^o * \nu^o$. Thus there exists a solution $\boldsymbol{P}^o$ such that each marginal distribution $\boldsymbol{P}^o(\eta_k \in \cdot)$ for any $k \in -\mathbf{N}$ coincides with ν^o. Thus we obtain (i).

Let $\boldsymbol{P}$ be a solution and let $\lambda = (\lambda_k : k \in -\mathbf{N})$ be the corresponding entrance law. Then we have

$$\lambda_k(\{g\}) = \frac{1}{2}\lambda_{k-1}(\{g\}) + \frac{1}{2}\lambda_{k-1}(\{g(12)\}), \quad g \in \mathfrak{S}_3. \tag{6.11}$$

This implies that there exist $p_0, p_1, p_2 \geq 0$ with $p_0 + p_1 + p_2 = 1$ such that

$$\lambda_k(\{e\}) = \lambda_k(\{(12)\}) = p_0/2, \tag{6.12}$$

$$\lambda_k(\{(13)\}) = \lambda_k(\{(23)\}) = p_1/2, \tag{6.13}$$

$$\lambda_k(\{(123)\}) = \lambda_k(\{(132)\}) = p_2/2 \tag{6.14}$$

for any $k \in -\mathbf{N}$. Therefore we obtain

$$\boldsymbol{P} = p_0\boldsymbol{P}^o + p_1\boldsymbol{P}^1 + p_2\boldsymbol{P}^2, \tag{6.15}$$

where $\boldsymbol{P}^1$ and $\boldsymbol{P}^2$ are defined in (6.9). Since the measures $\boldsymbol{P}^o$, $\boldsymbol{P}^1$ and $\boldsymbol{P}^2$ are mutually singular, we obtain (ii). Hence we obtain (iii). This completes the proof. □

7. Proof of the characterization theorem of the trichotomy

First, we prove Theorem 1.4. Before proving it, we need the following.

Lemma 7.1. *The set $\mathcal{H} = \mathcal{H}^s_\mu$ is a* submodule *of $\mathcal{G}$, i.e., the following statements hold:*
(0) If $\rho_1 \in \mathcal{H}$ and if ρ_2 is equivalent to ρ_1, then $\rho_2 \in \mathcal{H}$.
(i) If $\rho_1, \rho_2 \in \mathcal{H}$, then $\rho_1 \otimes \rho_2 \in \mathcal{H}$.
(ii) If $\rho_1, \rho_2 \in \mathcal{H}$, then $\rho_1 \oplus \rho_2 \in \mathcal{H}$.
(iii) If $\rho \in \mathcal{H}$, then $\bar{\rho} \in \mathcal{H}$. Here $\bar{\rho}$ denotes the complex conjugate representation.

This is obvious, so we omit the proof.

We utilize the following fact.

Lemma 7.2 (van Kampen [5])**.** *Let $\mathcal{H}$ be a submodule of $\mathcal{G}$. Suppose that*

$$\rho(g) = \mathrm{id} \text{ for every } \rho \in \mathcal{H} \implies g = e \; . \tag{7.1}$$

Then $\mathcal{H} = \mathcal{G}$.

This fact plays a key role in the proof of the *Tannaka duality* in the representation theory of compact groups. For the proof of Lemma 7.2, see, e.g., [11, Theorem 13.1] and [3, Lemma 17].

Now we proceed to prove Theorem 1.4.

Proof of Theorem 1.4. The claim (i) is obvious by definition of the isotropic subgroup $H_\mu(\boldsymbol{P}^o)$. The claim (iii) follows immediately from (i) and (ii). Thus we need only to prove the claim (ii).

1°). Suppose that the case (**C2**) occurs, i.e., that there exists a strong solution $\boldsymbol{P} \in \mathcal{P}_\mu$. Then it holds that $\mathcal{F}^\eta_{-\infty} \subset \mathcal{F}^\xi_{-\infty}$ under $\boldsymbol{P}$. Since $(\xi_k : k \in -\mathbf{N})$ is an independent sequence, Kolmogorov's 0-1 law holds so that $\mathcal{F}^\eta_{-\infty}$ is $\boldsymbol{P}$-trivial. Then Lemma 2.7 says that the solution $\boldsymbol{P}$ must be an extremal point: $\boldsymbol{P} \in \mathrm{ex}(\mathcal{P}_\mu)$. Since $\mathcal{H}^s_\mu(\boldsymbol{P}) = \mathcal{G}$, we obtain $H^s_\mu = \{e\}$.

Theorem 4.3 says that all the solutions of $\boldsymbol{P}' \in \mathrm{ex}(\mathcal{P}_\mu)$ are obtained by $\boldsymbol{P}' = T_g(\boldsymbol{P})$ for some $g \in G$. Then it is clear that the solution $\boldsymbol{P}'$ is also strong. Therefore we obtain the last claim: All extremal solutions are strong and the others non-strong.

2°). Suppose that $H^s_\mu = \{e\}$. Then we see that

$$\rho(g) = \mathrm{id} \text{ for every } \rho \in \mathcal{H}^s_\mu \implies g = e. \tag{7.2}$$

Applying this fact and Lemma 7.1 to Theorem 7.2, we conclude that $\mathcal{H}^s_\mu$ coincides with the whole $\mathcal{G}$. This shows that all the extremal points are strong and that the case (**C2**) occurs. □

Second, we prove the inclusions (5.9) in Theorem 5.1. This is an immediate consequence of the following proposition, which generalizes Proposition 2 of [15].

Proposition 7.3. *Let μ be given and let $\boldsymbol{P} \in \mathcal{P}_\mu$.*
(i) *If $\rho \in \mathcal{H}^2_\mu$, then*

$$\boldsymbol{E}\left[\rho(\eta_k) \mid \mathcal{F}^\eta_{-\infty} \vee \mathcal{F}^\xi_k\right] = \rho(\eta_k), \quad k \in -\mathbf{N}. \tag{7.3}$$

If, moreover, $\boldsymbol{P} \in \mathrm{ex}(\mathcal{P}_\mu)$, then $\rho(\eta_k)$ is $\mathcal{F}^\xi_k$-measurable $\boldsymbol{P}$-a.s. for any $k \in -\mathbf{N}$.
(ii) *If $\rho \notin \mathcal{H}^1_\mu$, then*

$$\boldsymbol{E}\left[\rho(\eta_k) \mid \mathcal{F}^\eta_{-\infty} \vee \mathcal{F}^\xi_k\right] = O, \quad k \in -\mathbf{N}. \tag{7.4}$$

If, moreover, $\boldsymbol{P} \in \mathrm{ex}(\mathcal{P}_\mu)$, then $\rho(\eta_k)$ is never *$\mathcal{F}^\xi_k$-measurable $\boldsymbol{P}$-a.s. for any $k \in -\mathbf{N}$.*

Proof. (i) To prove the claim, it suffices to show that (7.3) for arbitrary small $k \in -\mathbf{N}$.

Since $\rho \in \mathcal{H}^2_\mu$, it holds that

$$\lim_{k \to -\infty} \prod_{j=k}^{k_0} |\det R_j| > 0 \tag{7.5}$$

for arbitrary small $k_0 \in -\mathbf{N}$. Iterating the equation (1.4), we have

$$\rho(\eta_{k_0}) = \rho(\eta_{k_0-n})\Xi_n, \quad n \in \mathbf{N} \tag{7.6}$$

where

$$\Xi_n = \rho(\xi_{k_0-n+1})\rho(\xi_{k_0-n+2}) \cdots \rho(\xi_{k_0}), \quad n \in \mathbf{N}. \tag{7.7}$$

Since $\det \boldsymbol{E}[\Xi_n] \neq 0$ for $n \in \mathbf{N}$, we can define

$$\Phi_n = (\boldsymbol{E}[\Xi_n])^{-1} \Xi_n, \quad n \in \mathbf{N}. \tag{7.8}$$

Then the sequence $(\Phi_n : n \in \mathbf{N})$ constitutes a matrix-valued bounded $(\mathcal{E}^{k_0}_n)$-martingale[4] where

$$\mathcal{E}^{k_0}_n = \sigma(\xi_{k_0}, \xi_{k_0-1}, \ldots, \xi_{k_0-n+1}), \quad n \in \mathbf{N}. \tag{7.9}$$

[4] We mean that $\boldsymbol{E}\left[\Phi_{n+1} \mid \mathcal{E}^{k_0}_n\right] = \Phi_n$.

Therefore Φ_n converges to an $\mathcal{F}^{\xi}_{k_0}$-measurable $V^\rho \otimes V^\rho$-valued random element Φ_∞ almost surely. Since

$$|\det \Phi_n| = (\det \boldsymbol{E}[\Xi_n])^{-1}, \quad n \in \mathbf{N}, \tag{7.10}$$

we obtain $\det \Phi_\infty \neq 0$ almost surely. Taking subsequence if necessary, we see that

$$\Psi_n = \rho(\eta_{k_0-n})\boldsymbol{E}[\Xi_n] = \rho(\eta_{k_0})(\Phi_n)^{-1}. \tag{7.11}$$

converges to an $\mathcal{F}^{\eta}_{-\infty}$-measurable $V^\rho \otimes V^\rho$-valued random element Ψ_∞ almost surely. Therefore we conclude that $\rho(\eta_{k_0}) = \Psi_\infty \Phi_\infty$ is $\mathcal{F}^{\eta}_{-\infty} \vee \mathcal{F}^{\xi}_{k_0}$-measurable.

(ii) We can easily prove the claim by imitating the proof of Proposition 2 of [15]. So we omit the proof. □

Third, we prove the rest of Theorem 5.1.

Proof of Theorem 5.1. 1°). It is obvious by definition that H^1_μ, H^2_μ and H^s_μ are closed normal subgroups of G.

2°). Let $\boldsymbol{P}^o \in \mathrm{ex}(\mathcal{P}_\mu)$ be fixed. We seek an equivalent expression of the condition that $g \in H_\mu(\boldsymbol{P}^o)$. Note that $T_g(\boldsymbol{P}^o) = \boldsymbol{P}^o$ if and only if

$$\boldsymbol{P}^o(g \cdot \eta_k \in \cdot) = \boldsymbol{P}^o(\eta_k \in \cdot), \quad k \in -\mathbf{N}, \tag{7.12}$$

which is equivalent to

$$\rho(g)\boldsymbol{E}^o[\rho(\eta_k)] = \boldsymbol{E}^o[\rho(\eta_k)], \quad k \in -\mathbf{N},\ \rho \in \mathcal{G} \tag{7.13}$$

by the Peter–Weyl theorem.

3°). Suppose that $g \in H^1_\mu$. Let $\rho \notin \mathcal{H}^1_\mu$ and $\boldsymbol{P}^o \in \mathrm{ex}(\mathcal{P}_\mu)$. Noting that

$$\boldsymbol{E}^o[\rho(\eta_k) \mid \mathcal{F}^{\xi}_{k-n}] = \rho(\eta_{k-n}) R_{k-n+1} R_{k-n+1} \cdots R_k, \tag{7.14}$$

we have

$$\left\| \boldsymbol{E}^o[\rho(\eta_k) \mid \mathcal{F}^{\xi}_{k-n}] \right\| \le \| R_{k-n+1} R_{k-n+1} \cdots R_k \|. \tag{7.15}$$

Letting $n \to \infty$, we have

$$\boldsymbol{E}^o[\rho(\eta_k) \mid \mathcal{F}^{\xi}_{-\infty}] = O. \tag{7.16}$$

Since $\mathcal{F}^{\xi}_{-\infty}$ is $\boldsymbol{P}^o$-trivial, we obtain

$$\boldsymbol{E}^o[\rho(\eta_k)] = O, \tag{7.17}$$

and then we see that (7.13) holds, which proves that $H^1_\mu \subset H_\mu(\boldsymbol{P}^o)$ for any $\boldsymbol{P}^o \in \mathrm{ex}(\mathcal{P}_\mu)$.

4°). Suppose that $g \in H_\mu(\boldsymbol{P}^o)$. Let $\rho \in \mathcal{H}^s_\mu$. Since $T_g(\boldsymbol{P}^o) = \boldsymbol{P}^o$, we have the following identity between two joint laws on $G^{-\mathbf{N}} \times G^{-\mathbf{N}}$:

$$\boldsymbol{P}^o((g \cdot \eta, \xi) \in \cdot) = \boldsymbol{P}^o((\eta, \xi) \in \cdot). \tag{7.18}$$

Thus we have the following identity between two regular conditional distributions on $G^{-\mathbf{N}}$ given $\mathcal{F}^{\xi}$ (cf. Proof of Lemma 4.5):

$$\boldsymbol{P}^o_\xi(g\cdot\eta\in\cdot)=\boldsymbol{P}^o_\xi(\eta\in\cdot)\quad \mu-\text{a.s.} \tag{7.19}$$

Hence we have

$$\boldsymbol{E}^o_\xi[\rho(g)\rho(\eta)]=\boldsymbol{E}^o_\xi[\rho(\eta)]\quad \mu-\text{a.s.} \tag{7.20}$$

Since $\rho(\eta)$ is $\mathcal{F}^{\xi}$-measurable, we obtain $\rho(g)\rho(\eta)=\rho(\eta)$, μ-a.s., which implies that $\rho(g)=\text{id}$. Therefore we obtain the inclusion $H_\mu(\boldsymbol{P}^o)\subset H^s_\mu$.

5°). The inclusion $H^s_\mu\subset H^2_\mu$ follows from the inclusion $\mathcal{H}^s_\mu\supset\mathcal{H}^2_\mu$, which is assured by Proposition 7.3.

6°). If G is abelian, then all irreducible representations of G are one-dimensional, so we obtain $\mathcal{H}^1_\mu=\mathcal{H}^2_\mu$, which implies (5.11). □

Finally, we prove the following lemma, which assures the convergence of $r^2_k[\rho]$.

Lemma 7.4. *Let $\mu_1,\dots,\mu_n$ be probability measures on $\{x\in\mathbf{C}^n : |x|:=(|x_1|^2+\cdots+|x_n|^2)^{1/2}\le 1\}$ and set*

$$u_i=\int x\mu_i(dx),\quad i=1,2,\dots,n. \tag{7.21}$$

Then

$$|\det(u_1\cdots u_n)|\le 1. \tag{7.22}$$

Proof.[5] By Hadamard's inequality, we have

$$|\det(u_1,\dots,u_n)|\le\prod_i|u_i|. \tag{7.23}$$

By Jensen's inequality, we have

$$|u_i|\le\int|x|\mu_i(dx)\le 1,\quad i=1,2,\dots,n. \tag{7.24}$$

This completes the proof. □

8. Tsirelson–Yor equations

Let S be a Polish space and G a compact group. Let $\theta:S\to G$ be a measurable map and let $\psi:G\times S\to S$ and $\psi^{-1}:G\times S\to S$ be two measurable maps such that

$$\psi^{-1}(g,\psi(g,s))=\psi(g,\psi^{-1}(g,s))=s,\quad g\in G,\ s\in S. \tag{8.1}$$

[5] The authors are informed of this simple proof by Y. Takahashi.

We consider the following stochastic equation in discrete negative time:

$$\eta_k = \psi\left(\theta(\eta_{k-1}), \xi_k\right), \quad k \in -\mathbf{N}. \tag{1.9}$$

Moreover, we assume the following.

Assumption 8.1. The mappings ψ and θ commute in the sense that

$$\theta(\psi(g, x)) = g \cdot \theta(x), \quad g \in G. \tag{8.2}$$

Definition 8.2. Let S, G, ψ, ψ^{-1} and θ as above and assume that Assumption 8.1 is satisfied. Then the stochastic equation (1.9) is called a *Tsirelson–Yor equation*, which will be abbreviated by "TYE".

Following the case of STYE's, we introduce the notion of a solution as follows. Let $\eta = (\eta_k, k \in -\mathbf{N})$ denote the coordinate mapping process on $S^{-\mathbf{N}}$ and set

$$\xi_k = \psi^{-1}\left(\theta\left(\eta_{k-1}\right), \eta_k\right), \quad k \in -\mathbf{N}. \tag{8.3}$$

The filtrations $(\mathcal{F}^\eta_k)$ and $(\mathcal{F}^\xi_k)$ are defined in the same way.

Definition 8.3. Let $\mu = (\mu_k, k \in -\mathbf{N})$ be a sequence of Borel probability measures μ_k on S. A *solution* of the TYE (1.9) with the noise law μ is a probability measure $\boldsymbol{P}$ on $S^{-\mathbf{N}}$ such that the following two statements hold:
(i) ξ_k is independent of $\mathcal{F}^\eta_{k-1}$ under $\boldsymbol{P}$, for any $k \in -\mathbf{N}$.
(ii) ξ_k is distributed as μ_k under $\boldsymbol{P}$, for any $k \in -\mathbf{N}$.

The totality of solutions of the TYE (1.9) with the noise law μ will be denoted by $\mathcal{P}_\mu$.

We adopt the same notions of strong solutions and uniqueness in law as are defined in Definitions 2.4 and 2.5.

If $S = G$, $\psi(g, s)$ is the product and θ is the identity mapping, then the TYE (1.9) is exactly the STYE on G and all the notions of a solution, a strong solution and uniqueness in law coincide.

Consider a TYE (1.9) with a given noise law μ. Denote $\widehat{\eta}_k = \theta(\eta_k)$, $\widehat{\mu} = \mu \circ \theta^{-\mathbf{N}}$ and so on. Then Assumption 8.1 implies that

$$\widehat{\eta}_k = \widehat{\eta}_{k-1} \cdot \widehat{\xi}_k, \quad k \in -\mathbf{N}. \tag{8.4}$$

This is nothing but the STYE on G with the noise law $\widehat{\mu}$. We define

$$\widehat{\boldsymbol{P}} = \text{the law of } \widehat{\eta} \text{ on } G^{-\mathbf{N}} \text{ under } \boldsymbol{P}. \tag{8.5}$$

Then we obtain a mapping

$$\mathcal{P}_\mu \ni \boldsymbol{P} \mapsto \widehat{\boldsymbol{P}} \in \widehat{\mathcal{P}}_{\widehat{\mu}} \tag{8.6}$$

Here $\widehat{\mathcal{P}}_{\widehat{\mu}}$ denotes the set of solutions of the STYE (8.4) with the noise law $\widehat{\mu}$.

Let us give two examples.

Example 8.4. Let $S = \mathbf{R}$ and $G = \mathbf{T} = \mathbf{R}/\mathbf{Z} \simeq [0,1)$. For $g \in [0,1)$ and $s \in \mathbf{R}$, we set $\psi(g,s) = s+g$ and $\psi^{-1}(g,s) = s-g$. Set $\theta(s) = \alpha(s)$ for $s \in \mathbf{R}$. Then the TYE (1.9) coincides with the equation

$$\eta_k = \alpha(\eta_{k-1}) + \xi_k, \quad k \in -\mathbf{N}. \tag{1.3}$$

Then the equation for $\widehat{\eta_k} = \alpha(\eta_k)$ is the STYE on $\mathbf{T}$, which is actually (1.12).

Example 8.5. Let $S = \mathbf{R} \setminus \{0\}$ and $G = \mathbf{Z}/2\mathbf{Z} \simeq \{1,-1\}$. For $g = \pm 1$ and $s \in \mathbf{R} \setminus \{0\}$, we set $\psi(\pm 1, s) = \pm s$ and $\theta(s) = \operatorname{sgn}(s)$. Then the TYE (1.9) coincides with the equation

$$\eta_k = \operatorname{sgn}(\eta_{k-1}) \cdot \xi_k, \quad k \in -\mathbf{N}. \tag{1.13}$$

Then the equation for $\widehat{\eta_k} = \operatorname{sgn}(\eta_k)$ is the STYE on $\mathbf{Z}/2\mathbf{Z}$, which is actually (1.14).

Let $g \in G$. For the coordinate process $(\eta_k : k \in -\mathbf{N})$ on $S^{-\mathbf{N}}$ and the process $(\xi_k : \in -\mathbf{N})$ defined by (8.3), we set

$$\eta'_k := \psi(g \cdot \theta(\eta_{k-1}), \xi_k), \quad k \in -\mathbf{N}. \tag{8.7}$$

For $\boldsymbol{P} \in \mathcal{P}_\mu$ for some μ, we define $T_g(\boldsymbol{P})$ by the law of the process $(\eta'_k : k \in -\mathbf{N})$ under $\boldsymbol{P}$.

Now we have the following.

Proposition 8.6. *(i) The mapping* (8.6) *is bijective.*
(ii) $\boldsymbol{P} \in \mathcal{P}_\mu$ *is strong iff so is* $\widehat{\boldsymbol{P}} \in \widehat{\mathcal{P}}_{\widehat{\mu}}$.
(iii) $\boldsymbol{P} \in \mathcal{P}_\mu$ *is extremal iff so is* $\widehat{\boldsymbol{P}} \in \widehat{\mathcal{P}}_{\widehat{\mu}}$.
(iv) The family of mappings $(T_g : g \in G)$ *defines a group action on* $\mathcal{P}_\mu$ *and its restriction on* $\operatorname{ex}(\mathcal{P}_\mu) \cap \mathcal{P}_\mu$ *is transitive.*
(v) The case (**C1**), (**C2**) *or* (**C3**) *occurs for the TYE* (1.9) *with the noise law* μ *iff so does for the STYE on* G *with the noise law* $\widehat{\mu}$, *accordingly.*

Proof. Let $\boldsymbol{P}' \in \widehat{\mathcal{P}}_{\widehat{\mu}}$ be given. For any $k \in -\mathbf{N}$, we define a probability measure π_k on $(G \times S)^{\{k,\dots,0\}}$ in the following way: Let $(U_k, \xi_k, \xi_{k+1}, \dots, \xi_0)$ be a family of independent random variables such that U_k is a G-valued random variable distributed as $\boldsymbol{P}'(\eta_k \in \cdot)$ and ξ_j is an S-valued random variable distributed as μ_j for $j = k, \dots, 0$. Set $\eta_k = \psi(U_k, \xi_k)$,

$$\eta_j = \psi(\theta(\eta_{j-1}), \xi_j), \quad j = k+1, \dots, 0 \tag{8.8}$$

and

$$U_j = \theta(\eta_j), \quad j = k+1, \ldots, 0. \tag{8.9}$$

Then we define

$$\pi_k = \text{the law of } ((U_j, \eta_j) : j = k, k+1, \ldots, 0). \tag{8.10}$$

Thanks to the consistency assumption (3.4), we see that the family $\{\pi_k : k \in -\mathbf{N}\}$ satisfies Kolmogorov's consistency condition. Therefore Kolmogorov's extension theorem ensures the existence of a probability measure $\boldsymbol{Q}$ on $(G \times S)^{-\mathbf{N}}$ whose projection on $(G \times S)^{\{k,\ldots,0\}}$ coincides with π_k for $k \in -\mathbf{N}$. If we define $\boldsymbol{P}$ by the projection of $\boldsymbol{Q}$ on $S^{-\mathbf{N}}$, then we obtain $\widehat{\boldsymbol{P}} = \boldsymbol{P}'$. Therefore we conclude that the mapping (8.6) is surjective.

The rest of the claims are obvious, so we omit their proofs. □

References

[1] Émery, M. and Schachermayer, W., *A remark on Tsirelson's stochastic differential equation*, Sém. Prob., XXXIII, 291–303, Lecture Notes in Math., **1709**, Springer, Berlin, 1999.

[2] Ikeda, N. and Watanabe, S., *Stochastic Differential Equations and Diffusion Processes*, North Holland-Kodansha, Amsterdam and Tokyo, 1981.

[3] Joyal, A. and Street, R., *An introduction to Tannaka duality and quantum groups*, Category theory (Como, 1990), Lecture Notes in Math., **1488** (1991), 411–492.

[4] Kallsen, J., *A stochastic differential equation with a unique (up to indistinguishability) but not strong solution*, Sém. Prob., XXXIII, 315–326, Lecture Notes in Math., **1709**, Springer, Berlin, 1999.

[5] van Kampen, E. R., *Almost periodic functions and compact groups*, Ann. of Math., **37** (1936), No. 1, 78–91.

[6] Le Gall, J. F. and Yor, M., *Sur l'équation stochastique de Tsirelson*, Sem. Prob. XVII, Lecture Notes in Math., **986** (1983), 81–88.

[7] Liptser, R. S. and Shiryayev, A. N., *Statistics of random processes I*, General theory; translated by A. B. Aries, Applications of Mathematics, Vol. 5, Springer-Verlag, New York, 1977.

[8] Revuz, D. and Yor, M., *Continuous martingales and Brownian motion*, third edition, Springer-Verlag, Berlin, 1999.

[9] Stroock, D. W. and Yor, M., *On extremal solutions of martingale problems*, Ann. Sci. Ecole Norm. Sup. IV, Ser. **13** (1980), no. 1, 95–164.

[10] Sugiura, M., *Unitary representations and harmonic analysis, An introduction*, Kodansha Ltd., Tokyo, Halstead Press [John Wiley & Sons], New York-London-Sydney, 1975.

[11] Tannaka, T., *Sôtsui genri (Duality principle)* (Japanese), Iwanami shoten, 1966.

[12] Tsirelson, B., *An example of a stochastic differential equation having no strong solution*, Theory Probab. Appl. **20** (1975), 416–418; translated from Teor. Veroyatn. Primen. **20** (1975), 427–430.

[13] Uenishi, C., *On weak solutions of stochastic differential equations* (Japanese), Master Thesis, March 2004, Ritsumeikan University.

[14] Yamada, T. and Watanabe, S., *On the uniqueness of solutions of stochastic differential equations*, J. Math. Kyoto Univ., **11** (1971), 155–167.

[15] Yor, M., *Tsirel'son's equation in discrete time*, Probab. Theory Related Fields, **91** (1992), 135–152.

[16] Zvonkin, A. K. and Krylov, N. V., *Strong solutions of stochastic differential equations*, Selecta Math. Sovietica, **1** (1981), no. 1, 19–61; translated by A. B. Aries, Proceedings of the School and Seminar on the Theory of Random Processes (Druskininkai, 1974), Part II (Russian) (1975), 9–88.

Jirô Akahori, Department of Mathematical Sciences, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga, 525-8577, Japan
E-mail address: akahori@se.ritsumei.ac.jp

Chihiro Uenishi, Cadem Corporation, Ltd., Shin-Yokohama Office, Houei-Shin-Yokohama Bldg., 2-14-9, Shin-Yokohama, Kouhoku-Ku, Yokohama, 222-0033, Japan

Kouji Yano, Research Institute for Mathematical Sciences, Kyoto University, Kyoto, 606-8502, Japan
E-mail address: yano@kurims.kyoto-u.ac.jp