Unitary Brownian motions are linearizable

B. Tsirelson

MSRI, Berkeley, and School of Mathematics, Tel Aviv University,
Tel Aviv 69978, Israel, e-mail: tsirel@math.tau.ac.il

1991 Mathematical Subject classification: primary 58D20; secondary 22E65, 28C20, 46G12, 46L57, 60B11, 60J65, 81S25.

Brownian motions in the infinite-dimensional group of all unitary operators are studied under strong continuity assumption rather than norm continuity. Every such motion can be described in terms of a countable collection of independent one-dimensional Brownian motions. The proof involves continuous tensor products and continuous quantum measurements. A by-product: a Brownian motion in a separable F-space (not locally convex) is a Gaussian process.

Introduction

The most celebrated and useful random process surely is the standard Brownian motion in $\mathbb{R}$ (Wiener process). It is Markovian and Gaussian. Its increments are independent and stationary. Its continuous sample paths are bizarre, but many associated probability distributions are smooth, and connected by wonderful formulas. The multidimensional standard Brownian motion can produce a lot of random processes by means of stochastic differential equations. Especially, it can produce its close relatives, well-known during half a century, — Brownian motions in Lie groups and other topological groups.

A Brownian motion in a Lie group G could be defined constructively, by means of its generator, an invariant differential operator of second order on G, or descriptively, as a continuous G-valued random process with stationary independent increments. The former (constructive) definition stipulates smoothness of the generator; the latter (descriptive) definition does not. Are they equivalent? The question was asked by Seizo Itô, and answered in the positive by Kôsaku Yosida [35]; see also Kiyosi Itô [16], and the book by Henry McKean [24, Sect. 4.7].

Brownian motions in infinite-dimensional groups arise naturally from stochastic differential equations; see [21, Chap. 4] ("temporally homogeneous Brownian flows"), [23, Chap. VIII], [33, Chap. IV], [27, Chap. 6]. Does the constructive approach exhaust all possibilities allowed by the descriptive approach? For the group $G = \mathrm{Diff}^r(M)$ of all diffeomorphisms of a compact smooth manifold M, the question is answered in the positive by Baxendale [5]; roughly speaking, Brownian motions X in G correspond naturally to Brownian motions Y in the tangent space $\Delta = T_e G$ to G (at the unit e of G) via the stochastic differential equation $(dX) \cdot X^{-1} = dY$ (in the sense of Stratonovich). For many other groups, for example, the group of all homeomorphisms of a manifold, the very idea of $T_e G$ becomes too vague. Is there any constructive approach to Brownian motions in such groups?

Let us split the problem in two. First, we waive the relation $\Delta = T_e G$ between Δ and G. Instead, we take the Hilbert space $\Delta = l_2$ once and for all (any other reasonable choice is equivalent, see 1.9). We try to construct a Brownian motion (X, Y) in the direct product $G \times \Delta$ (the additive group of Δ is meant), whose first component X is the given Brownian motion in G, the second component Y is some Brownian motion in Δ, and the two components are perfectly correlated in the following sense: for each $t \in (0, \infty)$ there is a one-to-one correspondence between sample paths of X and Y on $[0, t]$; that is, the σ-field $\mathcal{F}_t^X$ generated by $\big(X(s)\big)_{s\in[0,t]}$ coincides with the σ-field $\mathcal{F}_t^Y$ generated by $\big(Y(s)\big)_{s\in[0,t]}$. We call such (X, Y) a linearization of X.

So, the first part of the problem is linearizability, that is, existence of a linearization. It is a well-defined question; either X is linearizable, or not. If X is linearizable, we face the second part of the problem: to find a reasonable interpretation of Δ as a kind of $T_e G$, and of the perfect correlation between X and Y as a kind of the stochastic differential equation $(dX) \cdot X^{-1} = dY$. (It is not a well-defined question, in contrast to linearizability.)

The main result of the present work states that every Brownian motion in the unitary group $\mathrm{U}(H)$ of the Hilbert space H (of countable dimension) is linearizable. The linearizability is evidently inherited by each group G that possesses a continuous one-one homomorphism to $\mathrm{U}(H)$ (in other words, a faithful unitary representation); for example, the group of all measure preserving transformations of $(0, 1)$, or the group of all diffeomorphisms of a manifold. For the group of all homeomorphisms, however, the question is still open.

It is meant that the group $\mathrm{U}(H)$ is equipped with the strong (or, equivalently, weak) operator topology, rather than the norm topology. The distinction may be illustrated by the following simple (commutative) example. Let B_k be independent standard Brownian motions in $\mathbb{R}$. Define a (random) unitary operator $X(t)$ by $X(t)e_k = \exp(ic_k B_k(t))e_k$, where (e_k) is an orthogonal basis (thus, $X(t)$ is diagonal in the basis), and (c_k) is a sequence of positive numbers. Then $X(\cdot)$ is norm continuous for $c_k = 1/\sqrt{\log k}$, but not for $c_k = 1$. In the strong operator topology, however, $X(\cdot)$ is continuous in any case (even for $c_k = 2^k$)! Most of well-known results [33], [27] assume norm continuity, or even stronger conditions.

The crucial idea of splitting the constructive approach in two separate problems (linearizability, and its interpretation) was suggested to me by the closely related idea of linearizability in the theory of continuous tensor products of probability spaces (Feldman [13], Tsirelson and Vershik [34]). I am indebted to Anatoly Vershik for drawing my attention to the theory in 1994.

The present paper is self-contained; intersections and parallels with related works are noted, but may be ignored by the reader. Sect. 1 formulates notions and results. Sect. 2 develops a new criterion of linearizability for continuous tensor products of probability spaces. Sect. 3 relates unitary Brownian motions to the theory of continuous quantum measurements (see Davies [11]). Sect. 4 establishes local finiteness of the corresponding quantum stochastic process, which implies linearizability. It is interesting to observe quantum probability helping to classical probability. Naturally, driving forces behind the matter should be more intelligible for readers acquainted with continuous tensor products of Hilbert spaces and probability spaces [2, 13, 3, 4, 34], continuous quantum measurements

[11, 22, 7, 1], the relation between the former and the latter [6], and quantum stochastic calculus [25, 26, 15].

Brownian motions in a linear topological space are evidently Gaussian, if the space possesses sufficiently many continuous linear functions. However, L_p for $0 < p < 1$ is an example of a separable F-space (complete metric vector space) that possesses no non-zero continuous linear functions. It appears that the new criterion of linearizability (Sect. 2) is fulfilled for all Brownian motions in all linear spaces (in fact, in all commutative groups), which is shown in Sect. 5. It bridges a gap between two definitions of Gaussian measures. (Sect. 5 depends on Sect. 2, but does not depend on "quantal" sections 3,4.)

The following definition, used throughout the paper, is borrowed from [17, 5].

DEFINITION. A Brownian motion in a topological group G is a continuous G-valued random process X such that

(a) $X(0) = e$ (the unit of G);

(b) increments on the left,

$$X(t_1),\; X(t_2)(X(t_1))^{-1},\; \dots,\; X(t_n)(X(t_{n-1}))^{-1}\,,$$

are independent whenever $0 \le t_1 \le \ldots \le t_n < \infty$;

(c) the distribution of $X(t)(X(s))^{-1}$ for $0 \le s \le t < \infty$ depends on $t - s$ only.

For example, any Brownian motion in $\mathbb{R}$ (the additive group of $\mathbb{R}$ is meant) is of the form $X(t) = vt + \sigma B(t)$ for some $v \in \mathbb{R}$, $\sigma \in [0, \infty)$; here $B(t)$ is the standard Brownian motion in $\mathbb{R}$. Any Brownian motion on the circe $\{z \in \mathbb{C} : |z| = 1\}$ is of the form $X(t) = \exp(ivt + i\sigma B(t))$. The pair $\big(\exp(ivt + i\sigma B(t)),\, B(t)\big)$ is a Brownian motion in the product (of the circle and the real line), and forms a linearization of $\exp(ivt + i\sigma B(t))$ (for $\sigma \ne 0$, of course). However, a slightly different terminology is used in the next section: a Brownian motion in a group is treated as a representation of a *noise* in the group; a linearization of the motion is treated as a linearization of the noise; and a linearization of a noise is treated as a faithful representation of the noise in a linear space.

1. The white noise versus black noises

Multiplication of operators (or composition of transformations) gives rise to one-parametric semigroups of operators (or transformations). Multiplication of measure spaces (or tensor multiplication of Hilbert spaces) should give rise to one-parametric semigroups of such spaces. The idea appeared repeatedly, but still, notions and terminology for spaces are far less standard than these for operators. "Continuous tensor product systems of Hilbert spaces" are defined by Arveson [3] in a framework different from that of a theory of "complete Boolean algebras of type 1 factors" by Araki and Woods [2]. "Factorized Hilbert spaces (and probability spaces) over a Boolean algebra" are defined by Tsirelson and Vershik [34] in a framework different from that of Arveson, and of a theory of "factored probability spaces, indexed by a Borel σ-field" by Feldman [13]. (The short list of approaches is in no way exhaustive.) Throughout the paper I restrict myself to the definition given below, and use the shortest term "noise" borrowed from quantum stochastic calculus.

1.1 Definition. A *noise* consists of a probability space $(\Omega, \mathcal{F}, P)$, a one-parametric group (T_t) of measure preserving transformations $T_t : \Omega \to \Omega$ for $t \in \mathbb{R}$, and a two-parametric family $(\mathcal{F}_{s,t})$ of sub-σ-fields $\mathcal{F}_{s,t} \subset \mathcal{F}$ for $-\infty < s \le t < \infty$, such that for all r, s, t

(a) T_t sends $\mathcal{F}_{r,s}$ onto $\mathcal{F}_{r+t,s+t}$ $(r \le s)$;
(b) $\mathcal{F}_{r,s}$ and $\mathcal{F}_{s,t}$ are independent $(r \le s \le t)$;
(c) $\mathcal{F}_{r,s}$ and $\mathcal{F}_{s,t}$, taken together, generate $\mathcal{F}_{r,t}$ $(r \le s \le t)$.

1.2 Note. Here and henceforth, each probability space is assumed to be a Lebesgue space (in the sense of Rokhlin, see [30]), and each σ-field contains all sets of zero probability.

A noise will be called *trivial,* if each $\mathcal{F}_{s,t}$ is trivial, that is, consists of sets of probability 0 or 1 only. Otherwise, each $\mathcal{F}_{s,t}$ with $s < t$ is non-atomic.

Remind that a metric (space) is called Polish, if it is complete and separable. A Polish group is a metrizable topological group G that possesses a Polish metric ρ satisfying the condition

$$\rho(x_n, y_n) \to 0 \quad \text{if and only if both } x_n y_n^{-1} \to e \text{ and } y_n^{-1} x_n \to e$$

for all $x_n, y_n \in G$; here e is the unit of G. The condition requires more than ρ to conform to the topology of G. Every metrizable topological group possesses a left-invariant metric ρ_{left} (Birkhoff, Kakutani, see [20]). Existence of a right-invariant metric follows: $\rho_{\text{right}}(x, y) = \rho_{\text{left}}(x^{-1}, y^{-1})$. In general, no metric is both left-invariant and right-invariant. Given a left-invariant metric ρ_{left} and a right-invariant metric ρ_{right}, we may take $\rho(x, y) = \rho_{\text{left}}(x, y) + \rho_{\text{right}}(x, y)$; we have $\rho_{\text{left}}(x_n, y_n) \to 0 \iff \rho_{\text{left}}(y_n^{-1} x_n, e) \to 0 \iff y_n^{-1} x_n \to e$; similarly, $\rho_{\text{right}}(x_n, y_n) \to 0 \iff \rho_{\text{right}}(x_n y_n^{-1}, e) \to 0 \iff x_n y_n^{-1} \to e$, and we get $\rho(x_n, y_n) \to 0 \iff (y_n^{-1} x_n \to e \text{ and } x_n y_n^{-1} \to e)$. Thus, (x_n) is a Cauchy sequence if and only if both $x_m^{-1} x_n \to e$ and $x_n x_m^{-1} \to e$ for $m, n \to \infty$, which does not depend on the choice of ρ_{left}, ρ_{right}. So, G is Polish if and only if it is complete in the metric $\rho_{\text{left}} + \rho_{\text{right}}$ and separable. See [20], [19, 6-O and 6-Q on pp. 210–213], [5, Sect. 2].

1.3 Definition. Let $(\Omega, \mathcal{F}, P)$, (T_t), $(\mathcal{F}_{s,t})$ form a noise, and G be a Polish group. A *representation* of the noise in the group is a two-parametric family of G-valued random variables $X_{s,t}$ for $-\infty < s \le t < \infty$, such that for all r, s, t

(a) T_t sends $X_{r,s}$ to $X_{r+t,s+t}$ $(r \le s)$,
(b) $X_{s,t}$ is measurable w.r.t. $\mathcal{F}_{s,t}$ $(s \le t)$,
(c) $X_{s,t} X_{r,s} = X_{r,t}$ $(s \le t)$,

and for each neighborhoood U of the unit element of the group

(d) $P(X_{0,t} \in U) \to 1$ for $t \to 0+$.

The representation is called *continuous,* if for every such U

(e) $\frac{1}{t} P(X_{0,t} \notin U) \to 0$ for $t \to 0+$.

The representation is called *faithful,* if for each $t \ge 0$

(f) the σ-field $\mathcal{F}_{0,t}$ is generated by the set $\{X_{0,s} : s \in [0, t]\}$ of random variables.

1.4 Note. A G-valued random variable is an equivalence class of measurable maps $\Omega \to G$, the equivalence being equality almost everywhere.

It is well-known [5, Th. 3(i)] that the "continuity condition" (e) is equivalent to continuity of $X_{s,t}$ in s and t for almost all $\omega \in \Omega$ (provided that functions are appropriately chosen within the equivalence classes $X_{s,t}$).

1.5 Definition. Let G_1, G_2 be Polish groups.

(a) G_1 is *Brown subordinate* to G_2, if every noise that has a faithful continuous representation in G_1, necessarily has a faithful continuous representation in G_2.

(b) G_1 and G_2 are *Brown equivalent,* if both G_1 is Brown subordinate to G_2, and G_2 is Brown subordinate to G_1.

The classical result mentioned in Introduction implies that every n-dimensional Lie group is Brown equivalent to the n-dimensional linear space. It is easy to see that an m-dimensional linear space is Brown subordinate to an n-dimensional linear space if and only if $m \le n$. The proof of the following (main) result is finished at the end of Sect. 4.

1.6 Theorem. The infinite-dimensional unitary group $\mathrm{U}(H)$ is Brown equivalent to the Hilbert space (that is, to the additive group of the Hilbert space with the norm topology).

1.7 Note. Throughout the paper all Hilbert spaces are assumed to be separable and (unless otherwise stated) of infinite dimension. One Hilbert space, denoted by H, is complex; it is used as a carrier of the unitary group $\mathrm{U}(H)$. Another Hilbert space, denoted by Δ, is real; it is used as a carrier of "linear" Brownian motions. The group $\mathrm{U}(H)$ is the multiplicative group of all unitary operators on the Hilbert space H. The strong and the weak operator topologies coincide on $\mathrm{U}(H)$; this is the topology $\mathrm{U}(H)$ is equipped with. It is metrizable. Here is an example of a right-invariant metric: $\rho_{\text{right}}(U_1, U_2) = \sum_k 2^{-k} \|U_1 e_k - U_2 e_k\|$; here $e_1, e_2, \dots$ are an orthonormal basis of the Hilbert space. A sequence of unitary operators can converge strongly to a non-invertible isometric operator, which means that ρ_{right} is not complete. However, the metric $\rho(U_1, U_2) = \rho_{\text{right}}(U_1, U_2) + \rho_{\text{right}}(U_1^{-1}, U_2^{-1})$ is complete, thus $\mathrm{U}(H)$ is a Polish group. (A proof, given in [14, "Weak topology" on pp. 61–64] for the group of all measure preserving transformations of $(0,1)$, needs only trivial adaptation to $\mathrm{U}(H)$.)

1.8 Theorem. Every commutative Polish group is Brown subordinate to the Hilbert space.

A separable F-space may be defined as a linear topological space whose additive group is a Polish group [29, 20]. Local convexity is not assumed. The following facts are evident for locally convex spaces, and well-known for a number of specific F-spaces (such as L_p for $p < 1$). The full generality is achieved now by means of the new approach presented here.

1.9 Corollary. All infinite-dimensional separable F-spaces are Brown equivalent.

1.10 Corollary *(preliminary formulation).* A Brownian motion in a separable F-space is a Gaussian process.

See Sect. 5 for the final formulation of 1.10 (given after discussing some definitions of Gaussian processes and measures), and for proofs of 1.8–1.10.

1.11 Conjecture. There is a Polish group not Brown subordinate to the Hilbert space.

If a noise has a faithful continuous representation in some Polish group, then all its representations (in all Polish groups) are necessarily continuous. That is a consequence of Meyer's theorem on predictability [10] ensuring continuity of all martingales in the corresponding filtration. We may avoid using any theory of martingales by means of an equivalent formulation. Remind that, given a noise and some $t > s > 0$, the orthogonal projection from $L_2(\Omega, \mathcal{F}_{0,t}, P)$ onto $L_2(\Omega, \mathcal{F}_{0,s}, P)$ is the conditional expectation, $f \mapsto \mathbb{E}\,(f|\mathcal{F}_{0,s})$. The function $\mathbb{E}\,(f|\mathcal{F}_{0,s})$ is defined up to a negligible set depending on s. We avoid the trouble of non-countable union of negligible sets by restricting ourselves to rational numbers s.

1.12 Definition. Let $(\Omega, \mathcal{F}, P)$, (T_t), $(\mathcal{F}_{s,t})$ form a noise. The noise is called *predictable,* if for any $t > 0$ and any $\mathcal{F}_{0,t}$-measurable bounded function $f : \Omega \to \mathbb{R}$, the function $\mathbb{E}\,(f|\mathcal{F}_{0,s})$, considered for rational $s \in (0,t)$, is uniformly continuous in s for almost all $\omega \in \Omega$.

The word "predictable", borrowed from the general theory of processes and filtrations, does not mean that the future of the noise can be predicted from its past. It means rather, that anything is predictable from the infinitesimally near past. Compare it with the Poisson process; its jumps are utterly unexpected, they have no precursors.

1.13 Lemma. If a noise has a faithful continuous representation in some Polish group then the noise is predictable.

1.14 Lemma. If a noise is predictable then all its representations in every Polish group are continuous.

Proofs of 1.13 and 1.14 are left to the reader.

Predictability is defined in a time-asymmetric way, since $\mathbb{E}\,(f|\mathcal{F}_{0,s})$ is considered rather than $\mathbb{E}\,(f|\mathcal{F}_{s,t})$. (You see, $f \ne \mathbb{E}\,(f|\mathcal{F}_{0,s}) + \mathbb{E}\,(f|\mathcal{F}_{s,t})$ in general.) Instead of time-reverse predictability, we may ask about predictability of time-reverse noise, formed by $(\Omega', \mathcal{F}', P') = (\Omega, \mathcal{F}, P)$, $T'_t = T_{-t}$, and $\mathcal{F}'_{s,t} = \mathcal{F}_{-t,-s}$. I do not know, whether predictability of a noise implies predictability of the time reverse noise, or not. Also, I do not know, whether every noise is isomorphic to its time reverse, or not. (A similar question is asked by Arveson [3, p. 6] about continuous tensor product systems of Hilbert spaces: "we do not know if an arbitrary product system must be antiisomorphic to itself".) Anyway, if a noise has a faithful continuous representation in some Polish group then, clearly, the time-reverse noise has a faithful continuous representation in the anti-isomorphic group (the same set G, but ba instead of ab); both noises are predictable, by Lemma 1.13.

If $(X_{s,t})$ is a representation of a predictable noise in (the additive group of) $\mathbb{R}$, then $(X_{0,t})_{t\in[0,\infty)}$ is a continuous process with stationary independent increments, that is, a Brownian motion in $\mathbb{R}$. All such representations are a linear space, that becomes a (real) Hilbert space H_{lin} of finite or countable dimension, being equipped with the norm $\|(X_{s,t})\| = \bigl(\mathbb{E}\,|X_{0,1}|^2\bigr)^{1/2}$. Trivial (non-random) representations $X_{s,t} = v(s-t)$, $v \in \mathbb{R}$, form a one-dimensional subspace; its orthogonal complement H^0_{lin} consists of centered (that is, zero-mean) representations; if $(X_{s,t}) \in H^0_{\text{lin}}$ and $\|(X_{s,t})\| = 1$, then $(X_{0,t})_{t\in[0,\infty)}$

is distributed as the standard Brownian motion (Wiener process). See [3, Sect. 5] for the "dimension of a product system" parallel to our $\dim(H^0_{\mathrm{lin}})$. We may choose an orthonormal basis $(X^k_{s,t})$, $1 \le k < \dim H_{\mathrm{lin}}$, of H^0_{lin}. If $\dim H^0_{\mathrm{lin}} = d < \infty$, we get a representation $X_{s,t} = \bigl(X^1_{s,t}, \dots, X^d_{s,t}\bigr)$ of the noise in $\mathbb{R}^d$; otherwise, if $\dim H^0_{\mathrm{lin}} = \infty$, we may construct a representation $X_{s,t} = \bigl(c_1 X^1_{s,t}, c_2 X^2_{s,t}, \dots\bigr)$ in the Hilbert space l_2, provided that we choose positive constants $c_1, c_2, \dots$ decreasing fast enough. In any case, there is a representation $(X_{s,t})$ in the Hilbert space, containing (in an evident sense) all representations in $\mathbb{R}$. For any $r \le t$ consider the σ-field $\mathcal{F}^{\mathrm{lin}}_{r,t}$ generated by the set $\{X_{r,s} | s \in [r,t]\}$ of random variables. (The set $\{X_{s,t} | s \in [r,t]\}$ gives the same.) Clearly, $(\Omega, \mathcal{F}, P)$, (T_t), $(\mathcal{F}^{\mathrm{lin}}_{s,t})$ form a noise, that may be called the *linearizable part* of the given predictable noise $(\Omega, \mathcal{F}, P)$, (T_t), $(\mathcal{F}_{s,t})$.

1.15 Definition. A predictable noise is called *linearizable,* if $\mathcal{F}^{\mathrm{lin}}_{s,t} = \mathcal{F}_{s,t}$ for all $s \le t$.

A predictable noise is linearizable if and only if it has a faithful (continuous) representation in the Hilbert space.

The following fact results from an example of Tsirelson and Vershik [34, Sect. 5].

1.16 Theorem. There is a nontrivial predictable noise with trivial linear part (that is, having only trivial representations in $\mathbb{R}$).

The parallel fact for continuous tensor product systems of Hilbert spaces results from examples of R.T. Powers, see [3, Remark 5.4] and [4, p. 12].

So, a linearizable predictable noise may be decomposed into a finite or countable set of independent copies of the white noise. The opposite extreme is a predictable noise with trivial linear part. Such a noise may be called a black noise. Indeed, terms "white noise" and "coloured noise" (these are not noises as defined here) imply that a noise manifests itself through linear sensors. For a black noise, however, the response of any linear sensor is zero!

What could be a physically reasonable nonlinear sensor able to sense a black noise? Maybe, a fluid could do it, which is hinted at by the following words by Shnirelman [32, p. 3] about a paradoxical motion of an ideal incompressible fluid: "[...] very strong external forces are present, but they are infinitely-fast oscillating in space, and therefore are indistinguishable from zero in the sense of distributions. [...] This is the fault of the sensors, not of the forces."

I do not know, how many nonisomorphic black noises exist, but I believe they are a continuum (accordingly, the section is entitled "the white noise versus black noises", not "versus the black noise"). An invariant, proposed in the next section, seems to be able to distinguish a continuum of nonisomorphic black noises. A similar question about continuous tensor product systems of Hilbert spaces is asked by Arveson [4, p. 12]: "It is expected that Σ is uncountable, but this has not been proved." Another question of [4, p. 12] can be answered (in the positive) by means of Theorem 1.16, namely the question, "is there a nontrivial E_0-semigroup α with the property that there is a nonzero unit $U = \{U_t : t \ge 0\}$ and such that every other unit V is related to U by a relation of the form $V_t = e^{i\lambda t} U_t$, $t \ge 0$, where λ is a complex number?" In our language, a "unit" (for a noise) is a representation in the multiplicative semigroup of complex numbers; all such representations are trivial for a black noise [34, Th. 1.7].

2. Spectral type of a noise

Given a noise $(\Omega, \mathcal{F}, P)$, (T_t), $(\mathcal{F}_{s,t})$, consider the spaces $L_2(\mathcal{F}_{s,t}) = L_2(\Omega, \mathcal{F}_{s,t}, P)$. If $r \le s \le t$ then (under the canonical identification) $L_2(\mathcal{F}_{r,s}) \otimes L_2(\mathcal{F}_{s,t}) = L_2(\mathcal{F}_{r,t})$, which follows from 1.1(b,c). At the same time $L_2(\mathcal{F}_{r,s})$ and $L_2(\mathcal{F}_{s,t})$ are linear subspaces of $L_2(\mathcal{F}_{r,t})$; note that $L_2(\mathcal{F}_{r,s}) \cap L_2(\mathcal{F}_{s,t})$ is the one-dimensional space of constants, and $L_2(\mathcal{F}_{r,s}) + L_2(\mathcal{F}_{s,t})$ is (in general) much smaller than $L_2(\mathcal{F}_{r,t})$. Introduce σ-fields $\mathcal{F}_{-\infty,t}$, $\mathcal{F}_{t,+\infty}$, $\mathcal{F}_{-\infty,+\infty}$ naturally (say, $\mathcal{F}_{-\infty,t}$ is generated by all $\mathcal{F}_{s,t}$ with $s \in (-\infty, t]$), then $\mathcal{F}_{s,t} = \mathcal{F}_{-\infty,t} \cap \mathcal{F}_{s,+\infty}$ and $L_2(\mathcal{F}_{s,t}) = L_2(\mathcal{F}_{-\infty,t}) \cap L_2(\mathcal{F}_{s,+\infty})$.

Denote by $E_{s,t}$ the orthogonal projection from $L_2(\mathcal{F}_{-\infty,+\infty})$ onto $L_2(\mathcal{F}_{s,t})$. It is the conditional expectation: $E_{s,t} f = \mathbb{E}\,(f | \mathcal{F}_{s,t})$ for $-\infty \le s \le t \le +\infty$. Operators $E_{s,t}$ commute with each other, and $E_{s,t} = E_{r,t} E_{s,u}$ whenever $-\infty \le r \le s \le t \le u \le +\infty$; in particular, $E_{s,t} = E_{-\infty,t} E_{s,+\infty}$.

More generally, given $-\infty \le s_1 < t_1 < \ldots < s_n < t_n \le +\infty$, we may consider the σ-field generated by σ-fields $\mathcal{F}_{s_1,t_1}, \ldots, \mathcal{F}_{s_n,t_n}$; denote it by $\mathcal{F}_A$ where $A = (s_1, t_1) \cup \ldots \cup (s_n, t_n)$ is an elementary set, that is, a finite union of intervals. We identify elementary sets that coincide up to a finite number of points; say, $(r, s) \cup (s, t)$ is identified with (r, t). More exactly, we define the elementary Boolean algebra $\mathcal{A}$ as the factoralgebra of the Boolean algebra generated by intervals, modulo the ideal of finite sets. However, it is usual to say "an elementary set $A \in \mathcal{A}$" instead of "an equivalence class $A \in \mathcal{A}$", like "a function $f \in L_2$" instead of "an equivalence class $f \in L_2$". For instance, we may say that the complement of (s, t) in $\mathcal{A}$ is $(-\infty, s) \cup (t, +\infty)$.

So, we have σ-fields $\mathcal{F}_A$ and Hilbert spaces $L_2(\mathcal{F}_A)$ for $A \in \mathcal{A}$, satisfying $L_2(\mathcal{F}_{A \cup B}) = L_2(\mathcal{F}_A) \otimes L_2(\mathcal{F}_B)$ whenever $A \cap B = \emptyset$ in $\mathcal{A}$. (In terms of [34, Def. 1.2] we have a measure factorization over $\mathcal{A}$.) Corresponding orthogonal projections E_A satisfy $E_A f = \mathbb{E}\,(f | \mathcal{F}_A)$ and $E_A E_B = E_B E_A = E_{A \cap B}$. However, in general $E_A + E_B \ne E_{A \cup B} + E_{A \cap B}$, that is, $(1 - E_A)(1 - E_B) \ne 1 - E_{A \cup B}$. The map $A \mapsto E_A$ is not a homomorphism of Boolean algebras.

Commuting projections E_A generate a commutative von Neumann algebra of operators on the Hilbert space $L_2(\mathcal{F}_{-\infty,+\infty})$; the space decomposes into a direct integral of Hilbert spaces over the spectrum of the algebra, see [12, Appendix A84]. The spectrum is a (Lebesgue, see 1.2) measure space (Z, Σ, μ) (though the measure is determined up to equivalence), and each E_A becomes (the multiplication by) the indicator function $1_{e(A)}$ of a set $e(A) \in \Sigma$ (or rather, an equivalence class). The relation $E_A E_B = E_{A \cap B}$ for operators turns into the relation $e(A) \cap e(B) = e(A \cap B)$ for sets. So, we have a map, preserving intersections (but not a homomorphism) $A \mapsto e(A)$ from the Boolean algebra $\mathcal{A}$ to the Boolean algebra $\Sigma \bmod 0$.

2.1 Lemma. For any noise, the σ-field $\mathcal{F}_{-\infty,+\infty}$ is generated by the union of σ-fields $\mathcal{F}_{(-\infty,-\varepsilon) \cup (\varepsilon,+\infty)}$ over all $\varepsilon > 0$.

Proof. In general, an increasing family of σ-fields has at most a countable set of discontinuities (jumps). For the family $(\mathcal{F}_{-\infty,t})_{t \in \mathbb{R}}$ the set of discontinuities must be shift-invariant due to stationarity, see 1.1(a). Therefore the set is empty; $\mathcal{F}_{-\infty,t}$ depends on t continuously. In particular, $\mathcal{F}_{-\infty,0}$ is generated by all $\mathcal{F}_{-\infty,-\varepsilon}$. Similarly, $\mathcal{F}_{0,+\infty}$ is generated by all $\mathcal{F}_{\varepsilon,+\infty}$. However, $\mathcal{F}_{-\infty,0}$ and $\mathcal{F}_{0,+\infty}$ together generate $\mathcal{F}_{-\infty,+\infty}$. ∎

2.2 Corollary. For every $t \in \mathbb{R}$

$$\mu\bigl(Z \setminus e((-\infty, t-\varepsilon) \cup (t+\varepsilon, +\infty))\bigr) \to 0 \quad \text{for } \varepsilon \to 0 \, .$$

Proof. $E_{(-\infty,-\varepsilon)\cup(\varepsilon,+\infty)} \to 1$ for $\varepsilon \to 0$ strongly on $L_2(\mathcal{F}_{(-\infty,+\infty)})$ by Lemma 2.1, which proves the case $t = 0$. The general case follows by stationarity. ∎

Sets $e(A)$ are determined $\bmod 0$; to avoid troubles, restrict ourselves to rational elementary sets A (I mean that their boundary points must be rational). For any $z \in Z$ consider all rational A such that $z \in e(A)$. Such A are a filter (within the "rational" subalgebra), since $e(A) \subset e(B)$ whenever $A \subset B$, and $e(A \cap B) = e(A) \cap e(B)$. Though, it may happen that $z \in e(\emptyset)$, thus "filter" must be understood here as "proper or unproper filter"; the unproper filter contains all A. For each rational t there is A of the filter, bounded away from t, due to Corollary 2.2. Consider the intersection of all A of the filter; boundary points of these A may be included or excluded arbitrarily since, being rational, they cannot belong to the intersection. Denote the intersection by $C(z)$; it is a closed set. For every t, $\mu\{z : t \in C(z)\} = 0$ due to Corollary 2.2. Therefore $C(z)$ is of zero Lebesgue measure (for almost all z). Also, $C(z)$ is bounded, since $E_{(-t,t)} \to 1$ strongly, hence $\mu\bigl(Z \setminus e((-t,t))\bigr) \to 0$ for $t \to \infty$. So, $C(z)$ is a nowhere dense compact set.

Sets $e(A)$ for rational $A \in \mathcal{A}$ separate points of Z, therefore $z \in Z$ is uniquely determined by the corresponding filter $\{A : z \in e(A)\}$. The filter, in its turn, is uniquely determined by the corresponding intersection $C(z)$; namely, a rational $A \in \mathcal{A}$ belongs to the filter if and only if $C(z)$ is contained in the interior of A. So, $z \mapsto C(z)$ is an injective map from Z to the set $\mathcal{C}$ of all compact subsets of $\mathbb{R}$. The set $\mathcal{C}$ becomes a Polish space, being equipped with the Hausdorff metric. The Borel structure corresponding to the metric is the same as the Borel structure generated by sets of the form $\{C \in \mathcal{C} : C \subset A\}$ for all rational $A \in \mathcal{A}$ (treated as open sets), therefore the map is measurable. We may identify each z with $C(z)$, Z with $C(Z) \subset \mathcal{C}$, μ with a measure on the Polish space $\mathcal{C}$ (note that $\mu(\mathcal{C} \setminus Z) = 0$), and Σ with the σ-field of μ-measurable subsets of $\mathcal{C}$, which gives the following result.

2.3 Theorem. Let $(\Omega, \mathcal{F}, P)$, (T_t), $(\mathcal{F}_{s,t})$ form a noise. Then there are: a probability measure μ on the space $\mathcal{C}$ (of all compact subsets of $\mathbb{R}$, including the empty set), satisfying the condition

$$\mu\{C \in \mathcal{C} : t \in C\} = 0 \quad \text{for all } t \in \mathbb{R} \, , \tag{a}$$

and a direct integral decomposition

$$L_2(\Omega, \mathcal{F}_{-\infty,+\infty}, P) = \int_{\mathcal{C}}^{\oplus} \hat{L}(C) \, d\mu(C) \tag{b}$$

into a measurable field of Hilbert spaces $\hat{L}(C)$ such that for every elementary set (that is, a finite union of intervals) $A \subset \mathbb{R}$ and every $f \in L_2(\Omega, \mathcal{F}_{-\infty,+\infty}, P)$, decomposed as $f = \int_{\mathcal{C}}^{\oplus} \hat{f}(C) \, d\mu(C)$, $\hat{f}(C) \in \hat{L}(C)$, the conditional expectation is decomposed as

$$\mathbb{E}\,(f | \mathcal{F}_A) = \int_{\mathcal{C}}^{\oplus} \hat{f}(C) 1_{e(A)}(C) \, d\mu(C) \, , \tag{c}$$

where $e(A) = \{C \in \mathcal{C} : C \subset A\}$.

The measure μ is determined by the noise up to equivalence.

2.4 DEFINITION. The equivalence class of measures μ on the space $\mathcal{C}$ (of all compact subsets of $\mathbb{R}$), appearing in Theorem 2.3, is called the *spectral type* of the noise. Each such μ is called a *spectral measure* of the noise.

Given an interval $(s,t) \subset \mathbb{R}$, we may use the same construction for decomposing $L_2(\mathcal{F}_{s,t})$ into a direct integral of spaces $\hat{L}_{s,t}(C)$ over the space $\mathcal{C}_{s.t}$ of all compact subsets of (s,t) (the subsets are bounded away from s and t). Given $r < s < t$, we get two decompositions of the same Hilbert space,

$$\int^{\oplus} \hat{L}_{r,t}(C)\, d\mu_{r,t}(C) = L_2(\mathcal{F}_{r,t}) = L_2(\mathcal{F}_{r,s}) \otimes L_2(\mathcal{F}_{s,t}) =$$
$$= \Big(\int^{\oplus} \hat{L}_{r,s}(C_1)\, d\mu_{r,s}(C_1) \Big) \otimes \Big(\int^{\oplus} \hat{L}_{s,t}(C_2)\, d\mu_{s,t}(C_2) \Big) =$$
$$= \int^{\oplus} \hat{L}_{r,s}(C_1) \otimes \hat{L}_{s,t}(C_2)\, d\mu_{r,s}(C_1) d\mu_{s,t}(C_2)\,,$$

which means that for $\mu_{r,t}$-almost all C

$$\hat{L}_{r,t}(C) = \hat{L}_{r,s}(C \cap (r,s)) \otimes \hat{L}_{s,t}(C \cap (s,t)) \tag{2.5}$$

(the case $s \in C$ may be neglected), and $\mu_{r,t} = \mu_{r,s} \otimes \mu_{s,t}$. However, measures $\mu_{r,s}$, $\mu_{s,t}$, $\mu_{r,t}$ are not canonical, they are determined up to equivalence; it is better to write

$$\mu_{r,t} \sim \mu_{r,s} \otimes \mu_{s,t}\,. \tag{2.6}$$

(In terms of [34] it is not a measure factorization but a measure type factorization; the distinction is essential, see the example at the end of Sect. 1(c) of [34]. Another example, closer to (2.6), is the random set $C = \{t : X(t) = 1\}$, where X is the standard Brownian motion in $\mathbb{R}$; here, $C \cap (r,s)$ and $C \cap (s,t)$ are dependent, but their dependence may be expressed by a positive density over the product of marginals.)

On the other hand, $L_2(\mathcal{F}_{r,s})$ (as well as $L_2(\mathcal{F}_{s,t})$) is a subspace of $L_2(\mathcal{F}_{r,t})$ (in terms of the tensor product, f is identified with $f \otimes 1$, where 1 is the constant function $1(\omega) = 1$ treated as a special element of $L_2(\mathcal{F}_{s,t})$), and the corresponding orthogonal projection $f \mapsto \mathbb{E}\left(f | \mathcal{F}_{r,s}\right)$ transforms $\int^{\oplus}_{\mathcal{C}_{r,t}} \hat{f}(C)\, d\mu_{r,t}(C)$ into $\int^{\oplus}_{\mathcal{C}_{r,s}} \hat{f}(C)\, d\mu_{r,s}(C)$; here $\mathcal{C}_{r,s}$ is treated as a subset (rather than a factor) of $\mathcal{C}_{r,t}$. Thus, $\mu_{r,s} \sim \mu_{r,t}\big|_{\mathcal{C}_{r,s}}$; of course, the restricted measure $\mu_{r,t}\big|_{\mathcal{C}_{r,s}}$ is defined by $(\mu_{r,t}\big|_{\mathcal{C}_{r,s}})(E) = \mu_{r,t}(E \cap \mathcal{C}_{r,s})$. Similarly,

$$\mu_{s,t} \sim \mu\big|_{\mathcal{C}_{s,t}}\,. \tag{2.7}$$

The spectral measure μ on $\mathcal{C}$ emerges as follows. Any $f \in L_2(\mathcal{F}_{-\infty,+\infty})$ determines a finite measure μ_f on $\mathcal{C}$ such that

$$\mu_f\{C \in \mathcal{C} : C \subset A\} = \|\mathbb{E}\,(f | \mathcal{F}_A)\|^2 \tag{2.8}$$

for all elementary sets $A \subset \mathbb{R}$. There is f such that for every $g \in L_2(\mathcal{F}_{-\infty,+\infty})$, the corresponding measure μ_g is absolutely continuous w.r.t. μ_f (in fact, a "generic" f satisfies the condition). For any such f we may take $\mu = \mu_f$. Of course, $\mu_f(E) = \int_E \|\hat f(C)\|^2 \, d\mu(C)$.

In particular, consider the white noise generated by the standard Brownian motion X in $\mathbb{R}$. Any $f \in L_2(\mathcal{F}_{-\infty,+\infty})$ decomposes into multiple Itô integrals, $f = \sum_n \int \dots \int \hat f_n(t_1,\dots,t_n) \, dX(t_1) \dots dX(t_n)$. For the function $\mathbb{E}\,(f|\mathcal{F}_A)$ the decomposition is the same, but restricted to $t_1,\dots,t_n$ belonging to A. The measure μ_f is concentrated on finite sets C, and its n-point part is $|\hat f_n(t_1,\dots,t_n)|^2 \, dt_1 \dots dt_n$. So, μ is concentrated on finite sets, and its n-point part may be chosen as the n-dimensional Lebesgue measure. Spaces $\hat L(C)$ are one-dimensional, as far as the Brownian motion X is one-dimensional; if it is d-dimensional, then $\dim \hat L(\{t_1,\dots,t_n\}) = d^n$. Note also that the empty set $C = \emptyset \in \mathcal{C}$ is an atom for μ, and $\hat L(\emptyset)$ is the one-dimensional space of constants, which holds for any noise. We see that the spectral decomposition of a noise is a generalization of Itô decomposition for the white noise.

Consider the set $\mathcal{C}_1 \subset \mathcal{C}$ of all single-element sets C, that is, $C_1 = \{\{t\} \big| t \in \mathbb{R}\}$. It may happen that $\mu(\mathcal{C}_1) > 0$; in that case we get a nontrivial linear subspace

$$\hat L_1 = \int_{\mathcal{C}_1}^{\oplus} \hat L(C) \, d\mu(C) = \int_{\mathbb{R}}^{\oplus} \hat L_1(t) \, dt \subset L_2(\mathcal{F}_{-\infty,+\infty}) \, ;$$

the former integral is the same as in Th. 2.3 but restricted to $\mathcal{C}_1 \subset \mathcal{C}$; it may be transferred to $\mathbb{R}$ by the one-one correspondence $\mathcal{C}_1 \ni C = \{t\} \leftrightarrow t \in \mathbb{R}$, giving the latter integral; the Lebesgue measure (dt) is used, since the transferred measure is shift-invariant up to equivalence. Otherwise (when $\mu(\mathcal{C}_1) = 0$), $\hat L_1$ contains only 0. Clearly, $f \in \hat L_1$ if and only if $\mu_f(\mathcal{C} \setminus \mathcal{C}_1) = 0$. Each $f \in \hat L_1$ gives raise to a family $(f_{s,t})$ of $f_{s,t} \in \hat L_1 \cap L_2(\mathcal{F}_{s,t})$ for $s \le t$ such that $\mathbb{E}\, f_{s,t} = 0$, and $f_{r,s} + f_{s,t} = f_{r,t}$ whenever $r \le s \le t$, and $f_{-\infty,+\infty} = f$. (In terms of [34] we have an additive integral, see Def. 1.3 there.) On the other hand, if $f \in L_2(\mathcal{F}_{-\infty,+\infty})$ belongs to $L_2(\mathcal{F}_{-\infty,t}) + L_2(\mathcal{F}_{t,+\infty})$ for every t, and $\mathbb{E}\, f = 0$, then $f \in \hat L_1$. (Proof: μ_f is concentrated on $\{C \in \mathcal{C} : \emptyset \ne C \subset (-\infty, t)\} \cup \{C \in \mathcal{C} : \emptyset \ne C \subset (t, +\infty)\}$ for every t; the intersection over all rational t gives $\mathcal{C}_1$.) Remind now the linear part $(\mathcal{F}^{\mathrm{lin}}_{s,t})$ of a predictable noise, defined in Sect. 1 (before 1.15).

2.9 Lemma. For every predictable noise, every $f \in \hat L_1$ is measurable w.r.t. $\mathcal{F}^{\mathrm{lin}}_{-\infty,+\infty}$.

Proof. The space $\hat L_1$ is invariant under the one-parameter unitary group (U_t) of time shifts, corresponding to the given group (T_t) of measure preserving transformations. Another one-parameter unitary group (V_λ) acting on $\hat L_1$ (but not the whole L_2) consists of diagonalizable operators (see [12, Appendix A80]) $V_\lambda = \int_{\mathbb{R}}^{\oplus} e^{i\lambda t} \, dt$ on $\hat L_1 = \int_{\mathbb{R}}^{\oplus} \hat L_1(t) \, dt$. The two groups satifsy Weil relation $V_\lambda U_t = e^{i\lambda t} U_t V_\lambda$. According to the von Neumann uniqueness theorem (see [28, Th. VIII.14 on p. 275]), $\hat L_1$ decomposes into direct sum of finite or countable number of irreducible components, — subspaces, each carrying an irreducible representation of $(U_t), (V_\lambda)$. Each irreducible representation is unitarily equivalent to the standard representation in $L_2(\mathbb{R})$, where U_t acts as the shift by t, and V_λ acts as the multiplication by $e^{i\lambda t}$. Comparing the latter with the formula $V_\lambda = \int_{\mathbb{R}}^{\oplus} e^{i\lambda t} \, dt$ we conclude that a function $f \in L_2(\mathbb{R})$ corresponds to $\int_{\mathbb{R}}^{\oplus} \hat h(t) f(t) \, dt$ for some vector field $\hat h(t) \in \hat L_1(t)$

(not depending on f). The irreducible component number k ($1 \le k < 1+d$, where $d \in \{0,1,\dots,\infty\}$ is the number of components) determines its vector field $\hat{h}_k(t) \in \hat{L}_1(t)$, and the set $\{\hat{h}_k(t) | 1 \le k < 1+d\}$ is an orthonormal basis of $\hat{L}_1(t)$. Comparing the action of U_t on $L_2(\mathbb{R})$ and $\hat{L}_1$ we conclude that $U_t \hat{h}_k(s) = \hat{h}_k(s+t)$. For each k we construct a representation $(X^k_{s,t})$ of the noise in $\mathbb{R}$ as follows: $X^k_{s,t} = \int^{\oplus}_{(s,t)} \hat{h}_k(u)\,du$. All $X^k_{s,t}$ are measurable w.r.t. $\mathcal{F}^{\mathrm{lin}}_{-\infty,+\infty}$. Every element of $\hat{L}_1$ is of the form $\sum_k \int^{\oplus}_{\mathbb{R}} \hat{h}_k(t) f_k(t)\,dt$, therefore it is also measurable w.r.t. $\mathcal{F}^{\mathrm{lin}}_{-\infty,+\infty}$. ∎

The same argument can be applied to a non-predictable noise, giving both Gaussian and Poissonian components of the linear part of the noise, but we do not need it.

Note that $\dim H^0_{\mathrm{lin}}$, discussed in Sect. 1 (before 1.15), is equal to $\dim \hat{L}_1(t)$, that is, $\dim \hat{L}(C)$ for $C \in \mathcal{C}_1$.

2.10 Corollary. For every predictable noise, if $\hat{L}_1$ generates the whole σ-field $\mathcal{F}_{-\infty,+\infty}$ then the noise is linearizable.

A related result about continuous tensor product systems of Hilbert spaces is given by Arveson [4, Theorem E in Sect. 6]. His "decomposable operators" correspond to a multiplicative counterpart of $\hat{L}_1$, — "multiplicative integrals" in terms of [34], while elements of $\hat{L}_1$ are "additive integrals". The two kinds of integrals generate the same σ-field [34, Th. 1.7].

2.11 Corollary. The following conditions are equivalent for every predictable noise.

(a) The linear part of the noise is trivial.

(b) $H^0_{\mathrm{lin}} = \{0\}$.

(c) $\hat{L}_1 = \{0\}$.

(d) $\mu(\mathcal{C}_1) = 0$.

(e) μ is concentrated on sets C with no isolated points.

Proof. (a) $\Longleftrightarrow$(b) by definitions; (b) $\Longleftrightarrow$(c) since $\dim H^0_{\mathrm{lin}} = \dim \hat{L}_1(t)$; (c) $\Longleftrightarrow$(d) by definition of $\hat{L}_1$; (e) $\Longrightarrow$(d) trivially; and (d) $\Longrightarrow$(e) due to (2.7). ∎

Turn to the set $\mathcal{C}_{\mathrm{finite}} \subset \mathcal{C}$ of all finite sets C (including the empty set). The corresponding subspace

$$\hat{L}_{\mathrm{finite}} = \int^{\oplus}_{\mathcal{C}_{\mathrm{finite}}} \hat{L}(C)\,d\mu(C) \subset L_2(\mathcal{F}_{-\infty,+\infty})$$

consists of all $f \in L_2(\mathcal{F}_{-\infty,+\infty})$ such that $\mu_f(\mathcal{C} \setminus \mathcal{C}_{\mathrm{finite}}) = 0$; the space is non-trivial if and only if $\mu(\mathcal{C}_{\mathrm{finite}}) > \mu(\{0\})$.

2.12 Theorem. $\hat{L}_{\mathrm{finite}} = L_2\big(\mathcal{F}^{\mathrm{lin}}_{-\infty,+\infty}\big)$ for every predictable noise.

Proof. $L_2\big(\mathcal{F}^{\mathrm{lin}}_{-\infty,+\infty}\big) \subset \hat{L}_{\mathrm{finite}}$ due to the decomposition into multiple Ito integrals, since the linearizable part of the noise is generated by independent Brownian motions in $\mathbb{R}$. In order to prove that $\hat{L}_{\mathrm{finite}} \subset L_2\big(\mathcal{F}^{\mathrm{lin}}_{-\infty,+\infty}\big)$ note that $\mathcal{C}_{\mathrm{finite}}$ is the union of sets $\mathcal{C}_{t_1,\dots t_n} \subset \mathcal{C}$ defined for rational $t_1,\dots,t_n$ such that $-\infty < t_1 < \ldots < t_n < +\infty$ as follows: $C \in \mathcal{C}_{t_1,\dots t_n}$ if and only if each one of the $n+1$ intervals $(-\infty,t_1),(t_1,t_2),\dots,(t_{n-1},t_n),(t_n,+\infty)$

contains no more than one point of C, and no one of the points $t_1, \dots, t_n$ belongs to C. Consider the subspace $\hat{L}_{t_1,\dots,t_n} = \int^{\oplus}_{\mathcal{C}_{t_1\dots t_n}} \hat{L}(C)\,d\mu(C) \subset \int^{\oplus}_{\mathcal{C}_{\text{finite}}} \hat{L}(C)\,d\mu(C) = \hat{L}_{\text{finite}}$. We have $\mathcal{C}_{t_1\dots t_n} = \big(\{\emptyset\} \cup (-\infty, t_1)\big) \times \big(\{\emptyset\} \cup (t_1, t_2)\big) \times \dots \times \big(\{\emptyset\} \cup (t_n, +\infty)\big)$, where points t are identified with single-point sets $\{t\} \in \mathcal{C}$. Thus, $\hat{L}_{t_1\dots t_n} = \big(\hat{L}(\emptyset) \oplus \int^{\oplus}_{(-\infty,t_1)} \hat{L}_1(t)\,dt\big) \otimes \big(\hat{L}(\emptyset) \oplus \int^{\oplus}_{(t_1,t_2)} \hat{L}_1(t)\,dt\big) \otimes \dots \otimes \big(\hat{L}(\emptyset) \oplus \int^{\oplus}_{(t_n,+\infty)} \hat{L}_1(t)\,dt\big)$; here $\hat{L}(\emptyset)$ is the one-dimensional space of constants. Combining it with Lemma 2.9 we conclude that all elements of $\hat{L}_{t_1,\dots,t_n}$ are measurable w.r.t. $\mathcal{F}^{\text{lin}}_{-\infty,+\infty}$. It remains to note that the union of all $\hat{L}_{t_1,\dots,t_n}$ is dense in $\hat{L}_{\text{finite}}$. ∎

2.13 Corollary. If $\hat{L}_{\text{finite}}$ generates the whole σ-field $\mathcal{F}_{-\infty,+\infty}$, then the noise (assumed to be predictable) is linearizable.

2.14 Corollary. The following conditions are equivalent for every predictable noise.

(a) A spectral measure is concentrated on finite sets.

(b) The noise is the product of a finite or countable set of independent copies of the white noise.

2.15 Note. Consider the least α such that a spectral measure is concentrated on sets of Hausdorff dimension $\le \alpha$. It is an invariant of a noise. Probably, the invariant takes on a continuum of values, which could distinguish a continuum of nonisomorphic black noises.

3. From unitary Brownian motions to quantum stochastic processes

Brownian motions in a Lie group G are described by their generators, right-invariant second-order differential operators on G. For an n-dimensional G, such an operator depends on $\frac{1}{2}n^2 + \frac{3}{2}n$ real parameters. The second-order part of the operator is given by a symmetric tensor (a_{ij}) of diffusion coefficients (at the unit of G), and the first-order part contains n more parameters (b_i), often called the drift vector (at the unit of G), though they do not form a vector.

Let $G = \mathrm{U}(d)$ be the group of all unitary $d \times d$ matrices. The map $A \mapsto e^{iA}$ from the linear space Δ of all Hermitian $d \times d$ matrices to $\mathrm{U}(d)$ is smooth, and smoothly invertible in a neighborhood of the origin, which gives a natural local coordinate system on G; note that $n = d^2$. Denote the inverse map by $U \mapsto -i \log U$. A Brownian motion X in G determines a diffusion process (not a Brownian motion) A in Δ, $A(t) = -i \log X(t)$, well-defined for small t (until leaving the neighborhood).

The space Δ is a Euclidean space, with the Hilbert-Schmidt scalar product $(A,B)_{\text{HS}} = \operatorname{tr}(AB^*) = \operatorname{tr}(AB)$. Infinitesimal characteristics (b_i), (a_{ij}) of X may be identified with the following linear and bilinear forms on Δ: for all $B, C \in \Delta$,

$$b(B) = \frac{d}{dt}\Big|_{t=0} \mathbb{E} \operatorname{tr}\big(A(t)B\big), \quad a(B,C) = \frac{d}{dt}\Big|_{t=0} \mathbb{E}\big(\operatorname{tr}(A(t)B) \cdot \operatorname{tr}(A(t)C)\big).$$

There is another interesting parametrization. Matrices $x_t = \mathbb{E}\,X(t)$ form a one-parametric semigroup, therefore $x_t = e^{tY}$ for some Y; the matrix Y represents the drift of X in the linear space of all matrices (while b represents the drift of X in the manifold of unitary matrices). In order to get parameters describing the spread of X, note that each

$U \in \mathrm{U}(d)$ determines an automorphism $M \mapsto UMU^*$ of the matrix algebra $M_d(\mathbb{C})$, and the automorphism is a quadratic function of U. We define $\mathcal{T}_t(M) = \mathbb{E}\,\big(X(t)MX^*(t)\big)$, which gives a one-parameter semigroup of linear maps $\mathcal{T}_t : M_d(\mathbb{C}) \to M_d(\mathbb{C})$; thus, $\mathcal{T}_t = e^{tZ}$ for some linear map $Z : M_d(\mathbb{C}) \to M_d(\mathbb{C})$. (Do not confuse $\mathcal{T}_t$ with the time shift T_t introduced by Def. 1.1.) The following lemma shows that the pair (Y, Z) determines a and b uniquely. However, the result will not be used, and its proof is relegated to Appendix.

3.1 Lemma. (a) The semigroup $(\mathcal{T}_t)$ determines uniquely the law of the Brownian motion $\big(\det X(t)\big)^{-1/d} X(t)$ in the group $\mathrm{SU}(d)$.

(b) The two semigroups (x_t), $(\mathcal{T}_t)$ determine uniquely the law of the Brownian motion $X(t)$ in the group $\mathrm{U}(d)$.

Turn to the infinite-dimensional case: $G = \mathrm{U}(H)$ is the group of all unitary operators in the Hilbert space (see 1.7). We do not know, how to generalize infinitesimal characteristics a and b for all Brownian motions X in G (some cases are investigated in [33,27]). In contrast, Y and Z have a straightforward generalization. Still, operators

$$x_t = \mathbb{E}\, X(t) \tag{3.2}$$

are well-defined, $\|x_t\| \le 1$, and form a strongly continuous semigroup. The general theory of operator semigroups ensures that the semigroup has its generator Y, densely defined, usually unbounded (for detail, see [11]). However, we do not need the generator; what we need is the very semigroup (x_t). Another semigroup $(\mathcal{T}_t)$ is formed by bounded linear operators $\mathcal{T}_t : V \to V$, where V is the Banach space of all Hermitian trace-class operators on the Hilbert space H (see [11]); note that H is complex, while V is real. Operators $\mathcal{T}_t$ are defined as before:

$$\mathcal{T}_t(\rho) = \mathbb{E}\,\big(X(t)\rho X^*(t)\big) \tag{3.3}$$

for $\rho \in V$; they form a strongly continuous semigroup, and $\|\mathcal{T}_t\| \le 1$. In fact, $(\mathcal{T}_t)$ is a special case of a so-called quantum dynamical semigroup (see [22]), and $X(\cdot)$ is one of so-called enravelings of $(\mathcal{T}_t)$ (see [7]; physicists are more interested in nonlinear enravelings). So, any Brownian motion X in $\mathrm{U}(H)$ determines two semigroups, (x_t) on H and $(\mathcal{T}_t)$ on V. I do not know, whether X is uniquely determined by the semigroups, or not.

Remind the spectral measure μ on $\mathcal{C}$, and $\mu_{s,t}$ on $\mathcal{C}_{s,t}$. Denote by Σ the σ-field of all μ-measurable subsets of $\mathcal{C}$, and by $\Sigma_{s,t}$ the σ-field of all $\mu_{s,t}$-measurable subsets of $\mathcal{C}_{s,t}$.

3.4 Theorem. For every Brownian motion in the unitary group $U(H)$, there is one and only one family $(\mathcal{E}_{s,t})_{-\infty<s<t<+\infty}$ satisfying the following conditions (a)–(f).

(a) For any $s < t$, $\mathcal{E}_{s,t}$ is a map, defined on the σ-field $\Sigma_{s,t}$, taking on values in the algebra of all bounded linear maps from V to itself, and $\mathcal{E}_{s,t}(E_1) = \mathcal{E}_{s,t}(E_2)$ whenever E_1, E_2 differ by a $\mu_{s,t}$-negligible set.

(b) The map $\mathcal{E}_{s,t}$ is an instrument, as defined in [11, Chap. 4, Def. 1.1]. It means, first, countable additivity: for any sequence $E_1, E_2, \dots$ of disjoint sets in $\Sigma_{s,t}$

$$\mathcal{E}_{s,t}(E_1 \cup E_2 \cup \dots)\rho = \mathcal{E}_{s,t}(E_1)\rho + \mathcal{E}_{s,t}(E_2)\rho + \dots \tag{b1}$$

where the sum is norm convergent for each $\rho \in V$; second, positivity:

$$\mathcal{E}_{s,t}(E)\rho \in V_+ \quad \text{for all } \rho \in V_+,\ E \in \Sigma_{s,t}\,; \tag{b2}$$

(here $V_+ = \{\rho \in V : \forall h \in H \ \ (\rho h, h) \ge 0\}$); and third, trace conservation:

$$\operatorname{tr}\big(\mathcal{E}_{s,t}(\mathcal{C}_{s,t})\rho\big) = \operatorname{tr}(\rho) \quad \text{for all } \rho \in V\,. \tag{b3}$$

(There is one more condition, complete positivity, see [11, Sect. 9.2]; in fact, it is also satisfied, which, however, will be neither proved nor used.)

(c) $\mathcal{E}_{r,t}(E_1 \times E_2) = \mathcal{E}_{s,t}(E_2)\mathcal{E}_{r,s}(E_1)$ whenever $r < s < t$, $E_1 \in \Sigma_{r,s}$, $E_2 \in \Sigma_{s,t}$; here $E_1 \times E_2$ means $\{C \in \mathcal{C}_{r,t} : C \cap (r,s) \in E_1 \text{ and } C \cap (s,t) \in E_2\}$.

(d) $\mathcal{E}_{r+t,s+t}(T_t^{-1}E) = \mathcal{E}_{r,s}(E)$ whenever $r < s$, $t \in \mathbb{R}$, $E \in \Sigma_{r,s}$.

(e) $\mathcal{E}_{0,t}(\mathcal{C}_{0,t})\rho \to \rho$ (in norm) when $t \to 0+$, for every $\rho \in V$.

(f) $\mathcal{E}_{s,t}(\mathcal{C}_{s,t}) = \mathcal{T}_{t-s}$, and $\mathcal{E}_{s,t}\big(\{\emptyset\}\big)\rho = x_{t-s}\rho x_{t-s}^*$ whenever $\rho \in V$ and $-\infty < s < t < +\infty$. (Here (x_t) and $(\mathcal{T}_t)$ are defined by (3.2), (3.3).)

Before starting the proof, remind of some well-known notions and facts. Let $\psi \in G \otimes H$, where G is another Hilbert space. For any $h \in H$ define $(\psi, h)_H \in G$ by the equality $\big((\psi,h)_H, g\big) = \big(\psi, g \otimes h\big)$ for all $g \in G$. For any bounded linear operator $U : H \to G \otimes H$ define a bounded linear operator $\mathcal{T}_U : V \to V$ as follows. Choose an orthonormal basis (g_k) in G, and define $A_1, A_2, \dots : H \to H$ by $Uh = \sum_k g_k \otimes A_k h$ for all $h \in H$, then $\mathcal{T}_U(\rho) = \sum_k A_k \rho A_k^*$. For the one-dimensional operator $\rho_h \in V$ defined by $\rho_h x = (x,h)h$ for all $x \in H$ we have $\big(\mathcal{T}_U(\rho_h)h_2, h_1\big) = \sum_k (A_k h, h_1)\overline{(A_k h, h_2)} = \big((Uh,h_1)_H, (Uh,h_2)_H\big)$ for all $h_1, h_2 \in H$, which shows that $\mathcal{T}_U$ does not depend on the choice of the basis (g_k), since linear combinations of ρ_h are dense in V.

Let F, G be Hilbert spaces, and $U_1 : H \to F \otimes H$, $U_2 : H \to G \otimes H$. Define $U : H \to F \otimes G \otimes H$ as $U = (1_F \otimes U_2)U_1$. If $U_1 h = \sum_k f_k \otimes A_k h$ and $U_2 h = \sum_l g_l \otimes B_l h$, then $Uh = \sum_k f_k \otimes U_2 A_k h = \sum_{k,l} f_k \otimes g_l \otimes B_l A_k h$. It is easy to see that $\mathcal{T}_U = \mathcal{T}_{U_2}\mathcal{T}_{U_1}$, that is, $\mathcal{T}_U(\rho) = \mathcal{T}_{U_2}\big(\mathcal{T}_{U_1}(\rho)\big)$ for all $\rho \in V$.

Let $U : H \to G \otimes H$ be an isometric operator, and G be a direct integral over some measure space, $G = \int_\Lambda^\oplus G(\lambda)\,d\mu(\lambda)$. For any μ-measurable set $E \subset \Lambda$ consider the projection $Q_E = \int_\Lambda^\oplus 1_E(\lambda)\,d\mu(\lambda)$ on G; that is, if $g = \int_\Lambda^\oplus g(\lambda)\,d\mu(\lambda)$, then $Q_E g = \int_E^\oplus g(\lambda)\,d\mu(\lambda)$. Define $\mathcal{E}(E) = \mathcal{T}_{(Q_E \otimes 1_H)U}$. Then $\mathcal{E}$ is an instrument.

Proof of Theorem 3.4. Uniqueness: (c) and (f) determine $\mathcal{E}_{s,t}\big(e_{s,t}(A)\big)$, where $e_{s,t}(A) = \{C \in \mathcal{C}_{s,t} : C \subset A\}$ and $A \subset \mathbb{R}$ is a finite union of intervals. Sets $e_{s,t}(A)$ generate $\Sigma_{s,t}$ (mod 0), and $e_{s,t}(A \cap B) = e_{s,t}(A) \cap e_{s,t}(B)$. Therefore, $\mathcal{E}_{s,t}$ is determined on the whole $\Sigma_{s,t}$.

Existence. Introduce an isometric operator $\tilde{X}_t : H \to L_2(\mathcal{F}_{0,t}, H)$ by $(\tilde{X}_t h)(\omega) = X(t,\omega)h$ for $h \in H$, $\omega \in \Omega$. Identifying $L_2(\mathcal{F}_{0,t}, H)$ with $L_2(\mathcal{F}_{0,t}) \otimes H = \big(\int_{\mathcal{C}_{0,t}}^\oplus \hat{L}_{0,t}(C)\,d\mu_{0,t}(C)\big) \otimes H$ we have an isometric operator $\tilde{X}_t : H \to \big(\int^\oplus \dots\big) \otimes H$. An instrument $\mathcal{E}_{0,t}$ appears, $\mathcal{E}_{0,t}(E) = \mathcal{T}_{(Q_E \otimes 1_H)\tilde{X}_t}$. The same construction for $X_{s,t} = X(t)(X(s))^{-1}$ instead of $X_{0,t} = X(t)$ gives $\mathcal{E}_{s,t}$. The construction is invariant under the time shift group (T_t). Thus, (a), (b), and (d) hold.

For any $h_1, h_2 \in H$ we have $(\tilde{X}_t h_2, h_1)_H = (X(t)h_2, h_1)$, since $(\tilde{X}_t h_2, g \otimes h_1) = \int_\Omega \big(X(t,\omega)h_2, g(\omega)h_1 \big)\, dP(\omega) = \int_\Omega \big(X(t,\omega)h_2, h_1 \big) g(\omega)\, dP(\omega)$ for all $g \in L_2(\mathcal{F}_{0,t})$.

For any $\psi \in L_2(\mathcal{F}_{0,t}) \otimes H$, $h \in H$, and $E \in \Sigma_{0,t}$ we have $\big((Q_E \otimes 1_H)\psi, h \big)_H = Q_E(\psi, h)_H$, since the two expressions are linear in ψ, and coincide on factorizable vectors; indeed, if $\psi = f \otimes h_1$, $f \in L_2(\mathcal{F}_{0,t})$, $h_1 \in H$, then both expressions turn into $(h_1, h) Q_E f$.

Take $\psi = \tilde{X}_t h$, then $(\psi, h_1)_H = \big(X(t)h, h_1 \big)$, thus, $\big((Q_E \otimes 1_H)\tilde{X}_t h, h_1 \big)_H = Q_E\big(X(t)h, h_1 \big)$. Applying the formula $\big(\mathcal{T}_U(\rho_h)h_2, h_1 \big) = \big((Uh, h_1)_H, (Uh, h_2)_H \big)$ for $U = (Q_E \otimes 1_H)\tilde{X}_t$, we get the matrix element

$$\big(\mathcal{E}_{0,t}(E)(\rho_h)h_2, h_1 \big) = \big(Q_E(X(t)h, h_1), Q_E(X(t)h, h_2) \big) \tag{3.5}$$

for all $h, h_1, h_2 \in H$, $E \in \Sigma_{0,t}$.

For proving (f), consider the two extreme cases, $E = \{\emptyset\}$ and $E = \mathcal{C}_{0,t}$. For $E = \mathcal{C}_{0,t}$ we have $Q_E = 1_{L_2(\mathcal{F}_{0,t})}$ and $\big(\mathcal{E}_{0,t}(\mathcal{C}_{0,t})(\rho_h)h_2, h_1 \big) = \big((X(t)h, h_1), (X(t)h, h_2) \big) = \mathbb{E}\, (X(t)h, h_1)\overline{(X(t)h, h_2)}$. On the other hand, $\big(X(t)\rho_h X^*(t)h_2, h_1 \big) = (X^*(t)h_2, h)(X(t)h, h_1) = (X(t)h, h_1)\overline{(X(t)h, h_2)}$, and we get $\big(\mathcal{E}_{0,t}(\mathcal{C}_{0,t})(\rho_h)h_2, h_1 \big) = \mathbb{E}\, (X(t)\rho_h X^*(t)h_2, h_1) = \big(\mathcal{T}_t(\rho_h)h_2, h_1 \big)$ by (3.3). It means that $\mathcal{E}_{0,t}(\mathcal{C}_{0,t}) = \mathcal{T}_t$. Similarly (or by (e)), $\mathcal{E}_{s,t}(\mathcal{C}_{s,t}) = \mathcal{T}_{t-s}$, which is the first claim of (f). For the other case, $E = \{\emptyset\}$, we have $(Q_{\{\emptyset\}}f)(\omega) = \mathbb{E}\, f$ for all $\omega \in \Omega$, $f \in L_2(\mathcal{F}_{0,t})$, that is, $Q_{\{\emptyset\}}f = (\mathbb{E}\, f)1_\Omega$, which is 2.3(c) for $A = (-\infty, 0) \cup (t, +\infty)$. Thus, $\big(\mathcal{E}_{0,t}(\{\emptyset\})(\rho_h)h_2, h_1 \big) = \big(1_\Omega \mathbb{E}\, (X(t)h, h_1), 1_\Omega \mathbb{E}\, (X(t)h, h_2) \big) = \mathbb{E}\, (X(t)h, h_1) \cdot \overline{\mathbb{E}\, (X(t)h, h_2)} = (x_t h, h_1)\overline{(x_t h, h_2)}$. On the other hand, $\big(x_t \rho_h x_t^* h_2, h_1 \big) = (x_t^* h_2, h)(x_t h, h_1) = (x_t h, h_1)\overline{(x_t h, h_2)}$, and we get $\mathcal{E}_{0,t}(\{\emptyset\})(\rho_h) = x_t \rho_h x_t^*$ for all $h \in H$. Therefore $\mathcal{E}_{0,t}(\{\emptyset\})(\rho) = x_t \rho x_t^*$ for all $\rho \in V$. Similarly (or by (d)), $\mathcal{E}_{s,t}\big(\{\emptyset\}\big)\rho = x_{t-s}\rho x_{t-s}^*$, which completes the proof of (f).

Item (e) is reduced by (f) to strong continuity of the semigroup $(\mathcal{T}_t)$, that is, $\mathcal{T}_t \rho \to \rho$ when $t \to 0$. It suffices to prove it for $\rho = \rho_h$, $h \in H$. We have $\mathcal{T}_t(\rho_h) = \mathbb{E}\, \big(X(t)\rho_h X^*(t) \big) = \mathbb{E}\, \rho_{X(t)h}$; however, $\|\rho_{X(t)h} - \rho_h\| \le 2\|X(t)h - h\| \to 0$ when $t \to 0$ for almost all ω, and never exceeds $2\|h\|$; therefore $\|\mathcal{T}_t(\rho_h) - \rho_h\| \le 2\mathbb{E}\, \|X(t)h - h\| \to 0$, which proves (e).

Before proving (c), note that for any $r < s < t$, any linear operator $U : H \to L_2(\mathcal{F}_{s,t}, H)$ and any vector-function $\psi \in L_2(\mathcal{F}_{r,s}, H)$, the vector-function $\xi = \big(1_{L_2(\mathcal{F}_{r,s})} \otimes U \big)\psi$ is given by $\xi(\omega) = \big(U(\psi(\omega)) \big)(\omega)$. Indeed, it suffices to consider a factorizable $\psi = f \otimes h$, that is, $\psi(\omega) = f(\omega)h$, $f \in L_2(\mathcal{F}_{r,s})$, $h \in H$. Then $\xi = f \otimes Uh$, that is, $\xi(\omega) = f(\omega) \cdot \big((Uh)(\omega) \big)$, while $U(\psi(\omega)) = U(f(\omega)h) = f(\omega) \cdot (Uh)$ and $\big(U(\psi(\omega)) \big)(\omega) = f(\omega) \cdot \big((Uh)(\omega) \big)$, too.

Now assume that $r < s < t$, $E_1 \in \Sigma_{r,s}$, $E_2 \in \Sigma_{s,t}$; we have to prove that $\mathcal{E}_{r,t}(E_1 \times E_2) = \mathcal{E}_{s,t}(E_2)\mathcal{E}_{r,s}(E_1)$. Similarly to $\tilde{X}_{0,t}$, introduce $\tilde{X}_{s,t} : H \to L_2(\mathcal{F}_{s,t}, H) = L_2(\mathcal{F}_{s,t}) \otimes H$ by $(\tilde{X}_{s,t}h)(\omega) = X_{s,t}(\omega)h$; the same for $\tilde{X}_{r,s}$ and $\tilde{X}_{r,t}$. Denote $F = L_2(\mathcal{F}_{r,s})$, $G = L_2(\mathcal{F}_{s,t})$, $U_1 = \big(Q_{E_1} \otimes 1_H \big)\tilde{X}_{r,s} : H \to F \otimes H$, $U_2 = \big(Q_{E_2} \otimes 1_H \big)\tilde{X}_{s,t} : H \to G \otimes H$, and $U = \big(Q_{E_1 \times E_2} \otimes 1_H \big)\tilde{X}_{r,t} : H \to F \otimes G \otimes H$, then (c) takes the form $\mathcal{T}_U = \mathcal{T}_{U_2}\mathcal{T}_{U_1}$. It suffices to check that $U = (1_F \otimes U_2)U_1$. A simpler equality $\tilde{X}_{r,t} = (1_F \otimes \tilde{X}_{s,t})\tilde{X}_{r,s}$ holds, since $\big((1_F \otimes \tilde{X}_{s,t})\tilde{X}_{r,s}h \big)(\omega) = \big(\tilde{X}_{s,t}((\tilde{X}_{r,s}h)(\omega)) \big)(\omega) = \big(\tilde{X}_{s,t}(X_{r,s}(\omega)h) \big)(\omega) = X_{s,t}(\omega)X_{r,s}(\omega)h = X_{r,t}(\omega)h = (\tilde{X}_{r,t}h)(\omega)$. Also, $Q_{E_1 \times E_2} = Q_{E_1} \otimes Q_{E_2}$. Thus, $U = \big(Q_{E_1} \otimes Q_{E_2} \otimes 1_H \big)\big(1_F \otimes \tilde{X}_{s,t} \big)\tilde{X}_{r,s}$, while $(1_F \otimes U_2)U_1 = \big(1_F \otimes (Q_{E_2} \otimes 1_H)\tilde{X}_{s,t} \big)\big(Q_{E_1} \otimes 1_H \big)\tilde{X}_{r,s}$. It remains to check that $\big(Q_{E_1} \otimes Q_{E_2} \otimes 1_H\big)\big(1_F \otimes \tilde X_{s,t}\big) = \big(1_F \otimes (Q_{E_2} \otimes 1_H)\tilde X_{s,t}\big)\big(Q_{E_1} \otimes 1_H\big)$. The two linear operators $F \otimes H \to F \otimes G \otimes H$ coincide on the whole $F \otimes H$, since they coincide on factorizable vectors; for any $f \in F$, $h \in H$ both operators transform $f \otimes h$ into $Q_{E_1} f \otimes (Q_{E_2} \otimes 1_H)\tilde X_{s,t} h$. So, (c) is verified. ∎

The measure μ_f on $\mathcal{C}_{0,t}$, defined by (2.8) for any $f \in L_2(\mathcal{F}_{0,t})$, may be written in terms of spectral projections Q_E as $\mu_f(E) = (Q_E f, f) = \|Q_E f\|^2$. Let f be a matrix element of $X(t)$, that is, $f = \big(X(t)h_2, h_1\big)$ for arbitrary $h_1, h_2 \in H$, $t \in [0,\infty)$. There is a simple relation between the scalar-valued measure μ_f and the operator-valued measure $\mathcal{E}_{0,t}$. The former is a matrix element of the latter:

$$\mu_f(E) = \big(\mathcal{E}_{0,t}(E)\rho_{h_2}, \rho_{h_1}\big)_{\mathrm{HS}} \quad \text{for } f = \big(X(t)h_2, h_1\big)\,; \tag{3.6}$$

here $(\cdot,\cdot)_{\mathrm{HS}}$ is the scalar product in the Hilbert space of all Hilbert-Schmidt operators in H, that is, $(\rho_1, \rho_2)_{\mathrm{HS}} = \operatorname{tr}(\rho_1\rho_2)$ for $\rho_1, \rho_2 \in V$. (Only trace-class Hermitian operators are really used here.) Taking into account that $(A, \rho_h)_{\mathrm{HS}} = (Ah, h)$ for every A, we can deduce (3.6) from (3.5) as follows: $\big(\mathcal{E}_{0,t}(E)\rho_{h_2}, \rho_{h_1}\big)_{\mathrm{HS}} = \big(\mathcal{E}_{0,t}(E)(\rho_{h_2})h_1, h_1\big) = \big(Q_E(X(t)h_2, h_1), Q_E(X(t)h_2, h_1)\big) = \|Q_E(X(t)h_2, h_1)\|^2 = \|Q_E f\|^2 = \mu_f(E)$.

Conditions (a)–(e) of Theorem 3.4 define a notion close to the notion of a quantum stochastic process as defined in [11, Sect. 5.2]. There are two distinctions. First, Davies stipulates "the value space X"; assuming that his X is a single point, we get rid of it. Second, Davies considers only *finite* sets C, rather than all compact sets. This is the point! It will be shown in the next section, that our operator-valued measure $\mathcal{E}_{s,t}$ is concentrated on $\mathcal{C}_{\mathrm{finite}} \cap \mathcal{C}_{s,t}$. Thus, the noise will appear to be linearizable, and $(\mathcal{E}_{s,t})$ will appear to be a quantum stochastic process in the sense of Davies. However, his "bounded interaction rate" condition [11, (2.9) in Sect. 5.2] does not hold in our framework; finiteness of sets C is ensured by a more subtle mechanism described in the next section. Note also that the quantum dynamical semigroup $(\mathcal{T}_t)$ is strongly continuous ($\mathcal{T}_t\rho \to \rho$) but in general not norm continuous; compare it with the following phrases of Lindblad: "[...] we have to assume that the semigroup is norm continuous [...] a condition which is not fulfilled in many applications. (We may hope that this restriction can be ultimately removed using more powerful mathematics.)" [22, p. 120].

4. A compactness argument

Let X be a Brownian motion in the unitary group $\mathrm{U}(H)$, and $(\mathcal{E}_{s,t})$ the corresponding family of instruments, given by Theorem 3.4. Let $\rho \in V_1^+$, that is, $\rho : H \to H$, $\rho \ge 0$, and $\operatorname{tr}(\rho) = 1$. A probability measure $\mu_{s,t}(\rho,\cdot)$ on $(\mathcal{C}_{s,t}, \Sigma_{s,t})$ arises as follows: $\mu_{s,t}(\rho, E) = \operatorname{tr}\big(\mathcal{E}_{s,t}(E)\rho\big)$ for $E \in \Sigma_{s,t}$. The measure depends linearly on the parameter ρ. The V-valued measure $E \mapsto \mathcal{E}_{s,t}(E)\rho$ has its V-valued density $p_{s,t}(\rho,\cdot)$ w.r.t. $\mu_{s,t}(\rho,\cdot)$; it is a $\mu_{s,t}(\rho,\cdot)$-measurable function $\mathcal{C}_{s,t} \ni C \mapsto p_{s,t}(\rho, C) \in V_1^+$ (determined uniquely $\mu_{s,t}(\rho,\cdot)$-almost everywhere) such that $\mathcal{E}_{s,t}(E)\rho = \int_E p_{s,t}(\rho, C)\,\mu_{s,t}(\rho, dC)$ for all $E \in \Sigma_{s,t}$. Its existence can be checked easily by considering scalar-valued measures $E \mapsto \big((\mathcal{E}_{s,t}(E)\rho)h_2, h_1\big)$ for $h_1, h_2 \in H$, and their densities (Radon-Nikodym derivatives) w.r.t. $\mu_{s,t}(\rho,\cdot)$.

Property 3.4(c), $\mathcal{E}_{r,t}(E_1 \times E_2) = \mathcal{E}_{s,t}(E_2)\mathcal{E}_{r,s}(E_1)$, may be reformulated in terms of p and μ as follows. Let $\rho_0 \in V_1^+$. Consider $\rho_1 = \mathcal{E}_{r,s}(E_1)\rho_0 =$

$\int_{E_1} p_{r,s}(\rho_0, C_1)\mu_{r,s}(\rho_0, dC_1)$. We have $\int_{E_1\times E_2} p_{r,t}(\rho_0, C)\mu_{r,t}(\rho_0, dC) = \mathcal{E}_{r,t}(E_1 \times E_2)\rho_0 = \mathcal{E}_{s,t}(E_2)\rho_1 = \int_{E_1} \mathcal{E}_{s,t}(E_2)p_{r,s}(\rho_0, C_1)\mu_{r,s}(\rho_0, dC_1) = \int_{E_1} \mu_{r,s}(\rho_0, dC_1) \int_{E_2} \mu_{s,t}\big(p_{r,s}(\rho_0, C_1), dC_2\big)p_{s,t}\big(p_{r,s}(\rho_0, C_1), C_2\big)$ for all $E_1 \in \Sigma_{r,s}$, $E_2 \in \Sigma_{s,t}$. Taking the trace we get

$$\mu_{r,t}(\rho_0, E_1 \times E_2) = \int_{E_1} \mu_{r,s}(\rho_0, dC_1)\mu_{s,t}\big(p_{r,s}(\rho_0, C_1), E_2\big) \tag{4.1}$$

for $E_1 \in \Sigma_{r,s}$, $E_2 \in \Sigma_{s,t}$, and then

$$p_{r,t}(\rho_0, C_1 \cup C_2) = p_{s,t}\big(p_{r,s}(\rho_0, C_1), C_2\big) \tag{4.2}$$

for almost all $C_1 \in \mathcal{C}_{r,s}$, $C_2 \in \mathcal{C}_{s,t}$. The property (4.1) can be generalized to arbitrary $E \in \Sigma_{r,t}$; denote $E(C_1) = \{C_2 \in \mathcal{C}_{s,t} : C_1 \cup C_2 \in E\}$ for $C_1 \in \mathcal{C}_{r,s}$, then

$$\mu_{r,t}(\rho_0, E) = \int_{\mathcal{C}_{r,s}} \mu_{r,s}(\rho_0, dC_1)\mu_{s,t}\big(p_{r,s}(\rho_0, C_1), E(C_1)\big)\,. \tag{4.3}$$

It holds by (4.1) for every product set $E = E_1 \times E_2$, by additivity for finite unions of product sets, and by continuity for all sets E. Our next step is to make explicit a Markov property implicit in (4.2), (4.3).

Introduce Borel spaces $\mathcal{Y}_t = \mathcal{C}_{-\infty,t} \otimes V_1^+$. Let $s < t$. The instrument $\mathcal{E}_{s,t}$ gives raise to a transition probability $P_{s,t}$, that is, a Borel function on $\mathcal{Y}_s$ whose values are probability measures on $\mathcal{Y}_t$. Before giving a formal definition, consider the idea. We have $y_0 \in \mathcal{Y}_s$, that is, $y_0 = (C_0, \rho_0)$, $C_0 \in \mathcal{C}_{-\infty,s}$ and $\rho_0 \in V_1^+$. The latter determines the corresponding probability distribution $\mu_{s,t}(\rho_0, \cdot)$ for a random element C of $\mathcal{C}_{s.t}$. Choose C at random; calculate the corresponding $\rho_1 = p_{s,t}(\rho_0, C) \in V_1^+$. We get a (random) $y_1 \in \mathcal{Y}_t$, namely, $y_1 = (C_1, \rho_1)$, where $C_1 = C_0 \cup C$. The distribution of y_1 is the measure $P_{s,t}(y_0)$ on $\mathcal{Y}_t$ that corresponds to $y_0 \in \mathcal{Y}_s$. Formally,

$$P_{s,t}\big((C_0, \rho_0), A\big) = \mu_{s,t}\big(\rho_0, \{C \in \mathcal{C}_{s,t} : (C_0 \cup C, p_{s,t}(\rho_0, C)) \in A\}\big) \tag{4.4}$$

for all $C_0 \in \mathcal{C}_{-\infty,s}$, $\rho_0 \in V_1^+$, and Borel sets $A \subset \mathcal{Y}_t$. The Markov property is

$$P_{r,t}(y_0, A) = \int_{\mathcal{Y}_s} P_{r,s}(y_0, dy_1)P_{s,t}(y_1, A) \tag{4.5}$$

for all $y_0 \in \mathcal{Y}_r$ and all Borel sets $A \subset \mathcal{Y}_t$. A proof follows. We have $y_0 = (C_0, \rho_0)$, $C_0 \in \mathcal{C}_{-\infty,r}$, $\rho_0 \in V_1^+$. Introduce $E = \{C \in \mathcal{C}_{r,t} : \big(C_0 \cup C, p_{r,t}(\rho_0, C)\big) \in A\}$, then $P_{r,t}(y_0, A) = \mu_{r,t}(\rho_0, E)$ by (4.4), and $\mu_{r,t}(\rho_0, E)$ is given by (4.3). The right-hand side of (4.5) is an integral over $y_1 \in \mathcal{Y}_s$, but all relevant y_1 are of the form $y_1 = \big(C_0 \cup C_1, p_{r,s}(\rho_0, C_1)\big)$, and the distribution $P_{r,s}(y_0, \cdot)$ for y_1 results from the distribution $\mu_{r,s}(\rho_0, \cdot)$ for C_1 by (4.4). Now (4.5) takes the form $\int_{\mathcal{C}_{r,s}} \mu_{r,s}(\rho_0, dC_1)\mu_{s,t}\big(p_{r,s}(\rho_0, C_1), E(C_1)\big) = \int_{\mathcal{C}_{r,s}} \mu_{r,s}(\rho_0, dC_1)P_{s,t}\big((C_0 \cup C_1, p_{r,s}(\rho_0, C_1)), A\big)$; it remains to note that $P_{s,t}\big((C_0 \cup C_1, p_{r,s}(\rho_0, C_1)), A\big) = \mu_{s,t}\big(p_{r,s}(\rho_0, C_1), E(C_1)\big)$ for $C_1 \subset \mathcal{C}_{r,s}$ by (4.4) since, using (4.2),

$\{C_2 \in \mathcal{C}_{s,t} : \big(C_0 \cup C_1 \cup C_2, p_{s,t}(p_{r,s}(\rho_0, C_1), C_2)\big) \in A\} = \{C_2 \in \mathcal{C}_{s,t} : \big(C_0 \cup C_1 \cup C_2, p_{r,t}(\rho_0, C_1 \cup C_2)\big) \in A\} = \{C_2 \in \mathcal{C}_{s,t} : C_1 \cup C_2 \in E\} = E(C_1)$.

The Markov property (4.5) allows us to introduce a Markov process (Y_t). However, having (for now) no information on its regularity, we restrict ourselves to rational values of t, thus avoiding the choice of a modification. Strong Markov property is not claimed, only the simple (fixed-time) Markov property. Given $\rho_0 \in V_1^+$, there is a Markov process (Y_t), defined for all rational $t \in [0, \infty)$, such that Y_t takes on values in $\mathcal{Y}_t = \mathcal{C}_{-\infty,t} \times V_1^+$, that is,

$$\begin{aligned} &Y_t = (C_t, \rho_t), \quad Y_0 = (\emptyset, \rho_0) \text{ with probability } 1, \\ &\mathbb{P}\,(Y_t \in A | Y_s) = P_{s,t}(Y_s, A) \text{ for } s < t, A \subset \mathcal{Y}_t. \end{aligned} \tag{4.6}$$

The specific form (4.4) of transition probabilities implies the following. First, $C_t \cap (-\infty, 0) = \emptyset$, and $C_t \cap (0, s)$ does not depend on t as far as $t \ge s \ge 0$. Second, $\rho_t = p_{s,t}\big(\rho_s, C_t \cap (s,t)\big)$ for $t > s \ge 0$ (it is meant that s, t are rational). Third, $\mathbb{P}\big(C_t \cap (s,t) \in E \big| C_s, \rho_s\big) = \mu_{s,t}(\rho_s, E) = \operatorname{tr}\big(\mathcal{E}_{s,t}(E)\rho_s\big)$ for $t > s \ge 0$ and a Borel set $E \in \mathcal{C}_{s,t}$ (note that the probability does not depend on C_s, and depends linearly on ρ_s). It follows that $\mathbb{E}\,\big(\rho_t \big| C_s, \rho_s\big) = \int_{\mathcal{C}_{s,t}} p_{s,t}(\rho_s, C)\mu_{s,t}(\rho_s, dC) = \mathcal{E}_{s,t}(\mathcal{C}_{s,t})\rho_s = \mathcal{T}_{t-s}(\rho_s)$ by 3.4(f), so,

$$\mathbb{E}\,\big(\rho_t \big| C_s, \rho_s\big) = \mathcal{T}_{t-s}(\rho_s) \tag{4.7}$$

for rational s, t such that $t \ge s \ge 0$. The other part of 3.4(f) gives

$$\mathbb{P}\,\big(C_t \cap (s,t) = \emptyset \big| C_s, \rho_s\big) = \operatorname{tr}\big(x_{t-s}\rho_s x^*_{t-s}\big)\,. \tag{4.8}$$

The space V has its dual space V^*, identified with the space of all Hermitian operators $A : H \to H$; the natural bilinear form is of course $A, \rho \mapsto \operatorname{tr}(A\rho)$. The semigroup $(\mathcal{T}_t)$ on V has its dual semigroup $\mathcal{T}_t^*$ on V^*; we have $\operatorname{tr}\big(\rho\mathcal{T}_t^*(A)\big) = \operatorname{tr}\big(\mathcal{T}_t(\rho)A\big) = \mathbb{E}\operatorname{tr}\big(X(t)\rho X^*(t)A\big) = \mathbb{E}\operatorname{tr}\big(\rho X^*(t)AX(t)\big)$, thus,

$$\mathcal{T}_t^*(A) = \mathbb{E}\,\big(X^*(t)AX(t)\big) \tag{4.9}$$

for $A \in V^*$ and $t \ge 0$. Note that V is also a subset of V^*. The following property is peculiar for quantum dynamical semigroups generated by unitary Brownian motions:

$$\text{If } A \in V \qquad \text{then } \mathcal{T}_t^*(A) \in V \quad \text{and } \operatorname{tr}\big(\mathcal{T}_t^*(A)\big) = \operatorname{tr}(A)\,, \tag{4.10}$$

which follows from (4.9), since $\operatorname{tr}\big(X^*(t)AX(t)\big) = \operatorname{tr}\big((X(t))^{-1}AX(t)\big) = \operatorname{tr}(A)$, and $A \ge 0$ implies $X^*(t)AX(t) \ge 0$.

4.11 Lemma. Let $Q_n \in V^*$ be finite-dimensional projections, and $\varepsilon_n \to 0$ positive numbers. Then the set $K = \{\rho \in V_1^+ : \operatorname{tr}(Q_n\rho) \ge 1 - \varepsilon_n \text{ for } n = 1, 2, \dots\}$ is compact.

Proof. For every n the set $\{Q_n\rho Q_n : \rho \in V_1^+\}$ is finite-dimensional and bounded. It suffices to prove that $\|\rho - Q_n\rho Q_n\| \le 2\sqrt{\varepsilon_n}$ for all $\rho \in K$. Each $\rho \in V_1^+$ may be represented as $\rho = \mathbb{E}\,\rho_h$ for some random vector $h \in H$, $\|h\| = 1$. We have $\|\rho_h - Q_n\rho_h Q_n\|^2 = \|\rho_h - \rho_{Q_n h}\|^2 \le 4\|h - Q_n h\|^2 = 4\big(1 - \|Q_n h\|^2\big) = 4\big(1 - \operatorname{tr}(Q_n\rho_h)\big)$, therefore $\|\rho - Q_n\rho Q_n\|^2 \le \big(\mathbb{E}\,\|\rho_h - Q_n\rho_h Q_n\|\big)^2 \le \mathbb{E}\,\|\rho_h - Q_n\rho_h Q_n\|^2 \le 4\big(1 - \mathbb{E}\operatorname{tr}(Q_n\rho_h)\big) = 4\big(1 - \operatorname{tr}(Q_n\rho)\big) \le 4\varepsilon_n$ for $\rho \in K$. ∎

4.12 Lemma. Let $S_n \subset V^*$ be compact sets of compact operators, and $\varepsilon_n \to 0$ positive numbers. Then the set $K = \{\rho \in V_1^+ : \max_{A\in S_n} \operatorname{tr}(A\rho) \ge 1 - \varepsilon_n \text{ for } n = 1, 2, \dots\}$ is compact.

Proof. Compactness of an operator $A \in S_n$ implies existence of a finite-dimensional projection $Q_n^{(A)}$ such that $A \le Q_n^{(A)} + \varepsilon_n \cdot 1_H$. Compactness of the set S_n allows to choose a single finite-dimensional projection Q_n such that $A \le Q_n + \varepsilon_n \cdot 1_H$ for all $A \in S_n$. Then $\operatorname{tr}(A\rho) \le \operatorname{tr}(Q_n\rho) + \varepsilon_n$ for all $A \in S_n$, therefore $\operatorname{tr}(Q_n\rho) \ge \max_{A\in S_n} \operatorname{tr}(A\rho) - \varepsilon_n \ge 1 - 2\varepsilon_n$ for all $\rho \in K$. Lemma 4.11 completes the proof. ∎

4.13 Lemma. For every $\rho_0 \in V_1^+$, $t \in (0,\infty)$, and $\varepsilon > 0$ there is a compact set $K \subset V_1^+$ such that $\mathbb{P}\big(\rho_s \in K \text{ for all rational } s \in [0,t]\big) \ge 1 - \varepsilon$, where ρ_s is defined by (4.6).

Proof. Let $Q \in V^*$ be a finite-dimensional projection; introduce a martingale $M(s) = \mathbb{E}\big(\operatorname{tr}(Q\rho_t)\big|C_s, \rho_s\big)$. Using (4.7), $M(s) = \operatorname{tr}\big(Q\mathcal{T}_{t-s}(\rho_s)\big) = \operatorname{tr}(Q_s\rho_s)$ where $Q_s = \mathcal{T}^*_{t-s}(Q)$. Note that Q_s is continuous in s (the same argument as in the proof of 3.4(e)), therefore $\{Q_s : s \in [0,t]\}$ is a compact set. Taking into account that $M(s) \le 1$ always, we have $\mathbb{P}\big(\inf_{s\in[0,t]} M(s) < 1 - \varepsilon\big) \le \frac{1}{c}(1 - M(0)) = \frac{1}{c}\big(1 - \operatorname{tr}(Q\mathcal{T}_t(\rho_0))\big)$.

Choose finite-dimensional projections $Q_1, Q_2, \dots$ such that $\operatorname{tr}(Q_n\mathcal{T}_t(\rho_0)) \to 1$. Introduce compact sets $S_n = \{\mathcal{T}^*_{t-s}(Q_n) : s \in [0,t]\} \subset V^*$; each $\mathcal{T}^*_{t-s}(Q_n)$ is a compact operator due to (4.9). Choose $c_n \to 0$, $c_n > 0$ such that $\delta_n \to 0$, where $\delta_n = \big(1 - \operatorname{tr}(Q_n\mathcal{T}_t(\rho_0))\big)/c_n$. Martingales $M_n(s) = \operatorname{tr}\big(\mathcal{T}^*_{t-s}(Q_n)\rho_s\big)$ satisfy $\mathbb{P}\big(\inf_{s\in[0,t]} M_n(s) < 1 - c_n\big) \le \delta_n$. Choosing a subsequence we can get $\sum \delta_n < \infty$; then $\mathbb{P}(E_n) \ge 1 - \varepsilon_n$, where $E_n = \{\omega : M_k(s) \ge 1 - c_k \text{ for all } s \in [0,t] \text{ and } k = n, n+1, \dots\}$ and $\varepsilon_n = \delta_n + \delta_{n+1} + \dots \to 0$. Within E_n we have $\operatorname{tr}\big(\mathcal{T}^*_{t-s}(Q_k)\rho_s\big) \ge 1 - c_k$, therefore $\max_{A\in S_k} \operatorname{tr}(A\rho_s) \ge 1 - c_k$ for $k = n, n+1, \dots$ and all rational $s \in [0,t]$. Lemma 4.12 ensures that the set $K_n = \{\rho \in V_1^+ : \max_{A\in S_k} \operatorname{tr}(A\rho) \ge 1 - c_k \text{ for } k = n, n+1, \dots\}$ is compact. So, $\mathbb{P}\big(\rho_s \in K_n \text{ for all rational } s \in [0,t]\big) \ge \mathbb{P}(E_n) \ge 1 - \varepsilon_n \to 1$. ∎

For a given $\rho_0 \in V_1^+$ construct compact subsets $K_1 \subset K_2 \subset \dots$ of V_1^+ such that $\mathbb{P}\big(\rho_t \in K_n \text{ for all rational } t \in [0,n]\big) \ge 1 - 2^{-n}$ for all n. Introduce first exit times $\tau_n = \inf\{t : \rho_t \notin K_n\}$; these are stopping (Markov) times. (They may be irrational, which is harmless, since we do not need ρ_{τ_n}.) We have

$$\tau_1 \le \tau_2 \le \dots; \quad \tau_n \to \infty; \quad \rho_t \in K_n \text{ for } t < \tau_n; \quad K_1, K_2, \dots \text{ are compact.} \tag{4.14}$$

Denote $f_r(\rho) = \operatorname{tr}\big(x_r\rho x_r^*\big) = \operatorname{tr}\big(x_r^* x_r\rho\big)$ for $\rho \in V_1^+$, $r \in [0,\infty)$, then $\mathbb{P}\big(C_{s+r} = C_s\big|C_s, \rho_s\big) = f_r(\rho_s)$. We have $f_r(\rho) \to 1$ for $r \to 0$, since $f_r(\rho_h) = \operatorname{tr}\big(\rho_{x_r h}\big) = \|x_r h\|^2 \to \|h\|^2$ (here ρ_h is the one-dimensional operator, $\rho_h x = (x,h)h$). Functions $f_r(\cdot)$ are in fact linear functionals on V of norm ≤ 1, therefore the convergence $f_r(\cdot) \to 1$ must be uniform on compact subsets of V_1^+. We may choose rational $r_n > 0$ such that $f_{r_n}(\rho) \ge 1/2$ for $\rho \in K_n$; so,

$$\mathbb{P}\left(C_{t+r_n} = C_t \middle| C_t, \rho_t\right) \ge \frac{1}{2} \quad \text{for } n = 1, 2, \dots \text{ and all rational } t \in [0, \tau_n)\,. \tag{4.15}$$

Consider the number $\operatorname{card}\left(C_{t+r_n} \cap (t, \tau_n)\right)$ (maybe, $+\infty$) of points in C_{t+r_n} belonging to the (maybe empty) interval (t, τ_n).

4.16 Lemma. For every $k, n = 1, 2, \dots$ and every rational $t \in [0, \infty)$

$$\mathbb{P}\left(\operatorname{card}\left(C_{t+r_n} \cap (t, \tau_n)\right) \ge k\right) \le 2^{-k}\,.$$

Proof. Fix n and t; take an integer m, divide the interval $(t, t + r_n)$ into m intervals $I_1, \dots, I_m$ of equal length, and consider the (random) number N_m of intervals I_i such that $I_i \cap C_{t+r_n} \cap (t, \tau_n) \ne \emptyset$. It suffices to prove that $\mathbb{P}(N_m \ge k) \le 2^{-k}$ for all m, since $N_m \to \operatorname{card}\left(C_{t+r_n} \cap (t, \tau_n)\right)$ for $m \to \infty$, and moreover, the convergence is monotone for $m = 1, 2, 4, 8, \dots$ Further, it suffices to prove that $\mathbb{P}\left(N_m \ge k + 1 \middle| N_m \ge k\right) \le 1/2$ for $k = 0, 1, 2, \dots$ However, this fact follows from (4.15) applied for $t, t + (1/m)r_n, t + (2/m)r_n, \dots, t + ((m-1)/m)r_n$ by the standard argument with a Markov time, which is legal, since the Markov time takes on only a finite number of (rational) values. ∎

So, $C_t \cap [0, \tau_n]$ has a finite intersection with every interval of length r_n, therefore $C_t \cap [0, \tau_n]$ is finite; however, $\tau_n \to \infty$, thus C_t is finite for all t almost sure, that is,

$$\mu_{0,t}(\rho_0, \mathcal{C}_{\text{finite}}) = 1\,. \tag{4.17}$$

We have $\operatorname{tr}\left(\mathcal{E}_{0,t}(\mathcal{C}_{\text{finite}})\rho_0\right) = 1$; $\operatorname{tr}\left(\mathcal{E}_{0,t}(\mathcal{C} \setminus \mathcal{C}_{\text{finite}})\rho_0\right) = 0$; by positivity (3.4b2), $\mathcal{E}_{0,t}(\mathcal{C} \setminus \mathcal{C}_{\text{finite}})\rho_0 = 0$. However, ρ_0 is arbitrary; so,

$$\mathcal{E}_{0,t}(\mathcal{C} \setminus \mathcal{C}_{\text{finite}}) = 0\,. \tag{4.18}$$

Combining it with (3.6) we get $\mu_f(\mathcal{C} \setminus \mathcal{C}_{\text{finite}}) = 0$ for each f of the form $f = \left(X(t)h_2, h_1\right)$. It means that all such f belong to $\hat{L}_{\text{finite}}$. By Corollary 2.13, the noise is linearizable. Thus, the main part of Theorem 1.6 is achieved: the infinite-dimensional unitary group is Brown subordinate to the Hilbert space. The converse holds by the two following facts.

4.19 Note. Let G_1, G_2 be Polish groups. If there exists a continuous one-one homomorphism $G_1 \to G_2$, then G_1 is Brownian subordinate to G_2. (The proof is immediate.)

4.20 Note. There exists a continuous one-one homomorphism from (the additive group of) the Hilbert space to the unitary group.

Proof. Let $\xi_1, \xi_2, \dots$ be i.i.d. $N(0,1)$ random variables. Each $c = (c_1, c_2, \dots) \in l_2$ determines a random variable $\exp\left(i \sum c_k \xi_k\right)$, and the corresponding multiplication operator $U(c)$ on the space of all square integrable random variables. The operator $U(c)$ is unitary, and the map $c \mapsto U(c)$ is a continuous one-one homomorphism. ∎

5. The commutative case

Proof of Theorem 1.8. Let G be a commutative Polish group, and X a Brownian motion in G. Due to Corollary 2.13 it suffices to prove that $\varphi\big(X(t)\big) \in \hat{L}_{\text{finite}}$ for every bounded Borel function $\varphi : G \to \mathbb{R}$ and $t \in [0,\infty)$. That is, we have to prove that $\mu_f(\mathcal{C}_{\text{finite}}) = 1$, where $f = \varphi\big(X(t)\big)$, $\|f\|_{L_2} = 1$.

We have $f = \varphi(X_{0,t/2} + X_{t/2,t}) = \varphi(X_{t/2,t} + X_{0,t/2})$ due to commutatitity of G. It follows that the joint probability distribution (under μ_f) of $C \cap \big(0, \frac{t}{2}\big)$ and $C \cap \big(\frac{t}{2}, t\big)$ is the same as for $C \cap \big(\frac{t}{2}, t\big)$ and $C \cap \big(0, \frac{t}{2}\big)$. That is, μ_f is invariant under the piecewise linear transformation $\alpha : (0,t) \to (0,t)$, $\alpha(s) = s + \frac{t}{2}$ for $s \in \big(0, \frac{t}{2}\big)$, $\alpha(s) = s - \frac{t}{2}$ for $s \in \big(\frac{t}{2}, t\big)$. Likewise, μ_f is invariant under the group of all invertible piecewise linear transformations of $(0,t)$ having derivative $= 1$ on each piece (the group acts naturally on $\mathcal{C}_{0,t}$ modulo negligible sets).

Take an integer n, divide the interval $(0,t)$ into n subintervals of length t/n, and consider the probability distribution (under μ_f) of the number k of subintervals that intersect C; denote the probabilities by $a_0, a_1, \dots, a_n$ ($a_0 + \dots + a_n = 1$). On the other hand, for any $\varepsilon \in (0,1)$ consider $\delta(\varepsilon) = \mu_f\{C : C \cap [0, \varepsilon t] \ne \emptyset\}$; we have $\delta(\varepsilon) \to 0$ for $\varepsilon \to 0$ due to 2.3(a). Given the number k (of subintervals that intersect C), all the $\binom{n}{k}$ possibilities are equiprobable due to the invariance property of μ_f, and the conditional probability of $\{C : C \cap [0, \frac{m}{n} t] = \emptyset\}$ is $\binom{n-m}{k} / \binom{n}{k}$. So,

$$\delta\Big(\frac{m}{n}\Big) = 1 - \sum_{k=0}^{n} a_k \frac{\binom{n-m}{k}}{\binom{n}{k}}$$

for $m = 1, \dots, n$. However, $\binom{n-m}{k} / \binom{n}{k} = \frac{n-m}{n} \cdot \frac{n-m-1}{n-1} \cdot \ldots \cdot \frac{n-m-k+1}{n-k+1} \le \big(\frac{n-m}{n}\big)^k$, therefore $\delta\big(\frac{m}{n}\big) \ge 1 - \sum a_k \big(\frac{n-m}{n}\big)^k = \sum a_k \big(1 - (\frac{n-m}{n})^k\big)$, and

$$a_k + \dots + a_n \le \frac{\delta(\frac{m}{n})}{1 - (\frac{n-m}{n})^k}$$

for all k and m; the left-hand side does not depend on m.

Take an $\varepsilon \in (0,1)$ and consider $n = 2, 4, 8, \dots$ together with m_n such that $m_n/n < \varepsilon$, $m_n/n \to \varepsilon$ (while k does not depend on n). Note that $a_k + \dots + a_n$ (which should be denoted rigorously by $a_k^{(n)} + \dots + a_n^{(n)}$) tends to $\mu_f\{C : \operatorname{card}(C) \ge k\}$. We get

$$\mu_f\{C : \operatorname{card}(C) \ge k\} \le \frac{\delta(\varepsilon)}{1 - (1-\varepsilon)^k}$$

for all $k = 1, 2, \dots$ and $\varepsilon \in (0,1)$; the left-hand side does not depend on ε. It remains to choose $\varepsilon_k \to 0$ such that $(1-\varepsilon_k)^k \le 1/2$, getting $\mu_f\{C : \operatorname{card}(C) \ge k\} \le 2\delta(\varepsilon_k) \to 0$, and so, $\mu_f(\mathcal{C}_{\text{finite}}) = 1$. ∎

Proof of Corollary 1.9. A separable F-space is Brown subordinate to the Hilbert space due to Theorem 1.8; it suffices to prove that the Hilbert space is Brown subordinate to every infinite-dimensional separable F-space. It follows immediately from Note 4.19 and the following fact.

5.1 NOTE. Let Δ be the Hilbert space, and F an infinite-dimensional F-space. Then there exists a continuous one-one linear operator $\Delta \to F$.

PROOF. Take $x_1, x_2, \ldots \in F$ spanning an infinite-dimensional subspace. Define $T : l_2 \to F$ by $T(\xi_1, \xi_2, \ldots) = \sum \xi_n c_n x_n$, where c_n tend to 0 fast enough, then T is continuous. Take $\Delta = l_2 \ominus T^{-1}(0)$ and restrict T to Δ. ∎

The following fact will not be used formally, but is worth to be mentioned. In particular, it shows that the proof of Corollary 1.10 is simpler than it seems; path-to-path correspondences used are in fact point-to-point.

5.2 LEMMA. Let G be a commutative Polish group, X a Brownian motion in G, Δ the Hilbert space, and (X, Y) a Brownian motion in $G \times \Delta$ such that $\mathcal{F}_t^X \subset \mathcal{F}_t^Y$ for all $t \in [0, \infty)$. Then for every $t \in [0, \infty)$, $X(t)$ is measurable w.r.t. the σ-field generated by $Y(t)$.

PROOF. It suffices to prove that the random variable $f = \varphi(X(t))$ is measurable w.r.t. the σ-field generated by $Y(t)$ for every bounded Borel function $\varphi : G \to \mathbb{R}$. We may assume that Δ is a space of sequences, and $Y(t) = \big(B_1(t), B_2(t), \ldots\big)$, where $B_1, B_2, \ldots$ are independent standard Brownian motions (in $\mathbb{R}$). Like every element of $L_2(\mathcal{F}_t^Y)$, f can be decomposed into multiple Itô integrals,

$$f = \sum_{n=0}^{\infty} \int \cdots \int \sum_{k_1, \ldots, k_n} \hat{f}_{k_1, \ldots, k_n}(t_1, \ldots, t_n) \, dB_{k_1}(t_1) \ldots dB_{k_n}(t_n) \, .$$

Remind the invariance argument used in the proof of Theorem 1.8: piecewise linear transformations of $(0, t)$ (with derivative $= 1$ on each piece) act on paths of (X, Y) by measure preserving transformations, leaving invariant $X(t)$ and f. Therefore, $\hat{f}_{k_1, \ldots, k_n}(t_1, \ldots, t_n)$ remains unchanged, when the transformation acts on each of $t_1, \ldots, t_n$. It means that $\hat{f}_{k_1, \ldots, k_n}(t_1, \ldots, t_n)$ does not depend on $t_1, \ldots, t_n$, it is a constant! Thus, the stochastic integral is a polynomial of $Y(t)$ only. ∎

Denote by γ the standard Gaussian measure on the space $\mathbb{R}^\infty$ of all sequences of reals; that is, γ is the joint distribution of a sequence of i.i.d. $N(0, 1)$ random variables. Given a Polish group G, we introduce the set $L_0(\gamma, G)$ of all equivalence classes of γ-measurable functions $X : \mathbb{R}^\infty \to G$, the equivalence being the equality γ-almost everywhere.

5.3 LEMMA. Let G be a commutative Polish group. For any $X \in L_0(\gamma, G)$ the following three properties are equivalent.

(a) There exists $Y \in L_0(\gamma, G)$ such that $X(u) + X(v) = Y\big(2^{-1/2}(u + v)\big)$ for $\gamma \otimes \gamma$-almost all pairs $(u, v) \in \mathbb{R}^\infty \times \mathbb{R}^\infty$.

(b) For each $a \in l_2$ there exists $g \in G$ such that $X(u + a) = X(u) + g$ for γ-almost all $u \in \mathbb{R}^\infty$.

(c) There exist a subgroup $E \subset \mathbb{R}^\infty$ of γ-full measure, a homomorphism $X_1 : E \to G$, and an element $g \in G$ such that $X(u) = X_1(u) + g$ for γ-almost all u.

PROOF. (c) $\Longrightarrow$(a): $X(u)+X(v)=X_1(u)+g+X_1(v)+g=X_1(u+v)+2g$; we let $Y(w)=X_1(2^{1/2}w)+2g$, taking into account that $2^{1/2}w\in E$ for γ-almost all w.

(a) $\Longrightarrow$(b): $X(u+a)+X(v-a)=Y\bigl(2^{-1/2}(u+v)\bigr)=X(u)+X(v)$, therefore $X(u+a)-X(u)=X(v)-X(v-a)$ for $\gamma\otimes\gamma$-almost all (u,v), which means that both functions are constant γ-almost everywhere.

(b) $\Longrightarrow$(c). The function $X_0:l_2\to G$ defined by $X(u+a)=X(u)+X_0(a)$ is a homomorphism. If $a_1,a_2,\ldots\in l_2$, $a_n\to 0$ weakly in l_2, then $X_0(a_n)\to 0_G$ for $n\to\infty$, since $X(\cdot+a_n)\to X(\cdot)$ in probability. Introduce projections $P_n:\mathbb{R}^\infty\to\mathbb{R}^n\subset l_2\subset\mathbb{R}^\infty$, $P_n(\xi_1,\xi_2,\ldots)=(\xi_1,\ldots,\xi_n)=(\xi_1,\ldots,\xi_n,0,0,\ldots)$, and functions $X_n=X-X_0\circ P_n$, that is, $X_n(\xi_1,\xi_2,\ldots)=X(\xi_1,\xi_2,\ldots)-X_0(\xi_1,\ldots,\xi_n,0,0,\ldots)$, then $X_n(u+a)=X_n(u)$ for all $a\in\mathbb{R}^n$, which means that $X_n(\xi_1,\xi_2,\ldots)$ depends only on $\xi_{n+1},\xi_{n+2},\ldots$

Functions of the form $\varphi_n\circ P_n$ (for all n and all measurable $\varphi_n:\mathbb{R}^n\to G$) are dense in $L_0(\gamma,G)$ (equipped with the convergence in probability). Choose φ_n such that $\varphi_n\circ P_n\to X$ in probability. We have $\varphi_n\circ P_n-X\to 0$ in probability. Choose $\varepsilon_n\to 0$ such that $\gamma\{u:\rho\bigl(0,\varphi_n(P_nu)-X(u)\bigr)\le\varepsilon_n\}\ge 1-\varepsilon_n$. The probability of the event $\rho\bigl(0,\varphi_n(P_nu)-X(u)\bigr)\le\varepsilon_n$ is the expectation of the conditional probability of the same event, given the first n coordinates $\xi_1,\ldots,\xi_n$ of $u=(\xi_1,\xi_2,\ldots)$. Given n, we may choose $\xi_1,\ldots,\xi_n$ such that the conditional probability at $(\xi_1,\ldots,\xi_n)$ is $\ge 1-\varepsilon_n$. However, $\varphi_n(P_nu)-X(u)=\varphi_n(P_nu)-X_0(P_nu)-X_n(u)$. Denote $g_n=\varphi_n(\xi_1,\ldots,\xi_n)-X_0(\xi_1,\ldots,\xi_n,0,0,\ldots)$, then $\gamma\{u:\rho\bigl(0,g_n-X_n(u)\bigr)\le\varepsilon_n\}\ge 1-\varepsilon_n$; the conditioning is omitted, since $X_n(\xi_1,\xi_2,\ldots)$ depends only on $\xi_{n+1},\xi_{n+2},\ldots$ So, there exist $g_1,g_2,\ldots\in G$ such that $X_n-g_n\to 0$ in probability, which means that $X-X_0\circ P_n-g_n\to 0$, that is, $X_0\circ P_n+g_n\to X$ in probability, when $n\to\infty$. The following trick shows that $g_n\to g$ for some g.

Consider three measure preserving maps $\alpha,\alpha_1,\alpha_2:\bigl(\mathbb{R}^\infty\times\mathbb{R}^\infty,\gamma\otimes\gamma\bigr)\to(\mathbb{R}^\infty,\gamma)$ defined by $\alpha_1(u,v)=\frac12u+\frac{\sqrt3}{2}v$, $\alpha_2(u,v)=\frac12u-\frac{\sqrt3}{2}v$, $\alpha(u,v)=u$, then $\alpha_1(u,v)+\alpha_2(u,v)=\alpha(u,v)$. For any $a,b\in l_2$ we have $(X\circ\alpha_1+X\circ\alpha_2-X\circ\alpha)(u+a,v+b)-(X\circ\alpha_1+X\circ\alpha_2-X\circ\alpha)(u,v)=X_0(\alpha_1(a,b))+X_0(\alpha_2(a,b))-X_0(\alpha(a,b))=X_0\bigl(\alpha_1(a,b)+\alpha_2(a,b)-\alpha(a,b)\bigr)=0$, which means that the function $X\circ\alpha_1+X\circ\alpha_2-X\circ\alpha$ is constant; $X(\alpha_1(u,v))+X(\alpha_2(u,v))-X(\alpha(u,v))=g$ for $\gamma\otimes\gamma$-almost all (u,v). However, $X\circ\alpha_1=\lim_n(X_0\circ P_n\circ\alpha_1+g_n)$, and the same for α_2,α. We have $g=\lim_n(X_0\circ P_n\circ\alpha_1+g_n+X_0\circ P_n\circ\alpha_2+g_n-X_0\circ P_n\circ\alpha-g_n)=\lim_n(X_0\circ P_n\circ(\alpha_1+\alpha_2-\alpha)+g_n)=\lim_n g_n$.

So, $X_0\circ P_n+g\to X$ in probability, when $n\to\infty$. We choose $n_k\to\infty$ such that $X_0\circ P_{n_k}+g\to X$ almost sure. The set of all u such that $\lim_k X_0\bigl(P_{n_k}(u)\bigr)$ exists, is a subgroup E of $\mathbb{R}^\infty$, $\gamma(E)=1$, and the limit is the needed homomorphism $X_1:E\to G$. ∎

The convergence $X_0\circ P_n+g\to X$, obtained in the proof above, may be compared with other results [18, 31, 9]. There, convergence almost sure is established for linear spaces; here — convergence in probability, for commutative groups. Also, Lemma 5.3 may be compared with the study of "quasi-additive functionals" in [9]. There, maps $G\to\mathbb{R}$ are considered; here — maps $\mathbb{R}^\infty\to G$.

5.4 COROLLARY. Let G be (the additive group of) a separable F-space. Then, in Condition (c) of Lemma 5.3, the subgroup E can be chosen to be a linear subspace, and the homomorphism X_1 — a linear map.

Proof. The function $X_0 : l_2 \to G$ is linear, since it is a continuous homomorphism between F-spaces. Functions $X_0 \circ P_{n_k}$ are linear, therefore their limit X_1 is linear, and its domain E is a linear subspace. ∎

All reasonable definitions of a Gaussian measure on a separable Banach space are evidently equivalent, which cannot be said about separable F-spaces and Polish groups (see [9]).

A symmetric Gaussian measure in the sense of Fernique is a probability measure μ on a separable F-space F such that the product measure $\mu \otimes \mu$ on $F \oplus F$ is invariant under the following group of transformations: $T_\varphi(x,y) = \big(x \cos\varphi - y \sin\varphi, \, x \sin\varphi + y \cos\varphi \big)$ for $x, y \in F$, $\varphi \in \mathbb{R}$.

A Gaussian measure in the sense of Bernstein is a probability measure μ on a commutative Polish group G such that the product measure $\mu \otimes \mu$ on $G \times G$ turns into some product measure $\nu_1 \otimes \nu_2$ on $G \times G$ under the following transformation: $(x,y) \mapsto (x-y, x+y)$ for $x, y \in G$.

If μ is a symmetric Gaussian measure in the sense of Fernique, then μ (as well as any shift of μ) is also Gaussian in the sense of Bernstein.

A Gaussian (convolution) semigroup is a family $(\mu_t)_{t \in [0,\infty)}$ of probability measures μ_t on a Polish group G, produced by some Brownian motion $X(\cdot)$ in G, in the sense that μ_t is the distribution of $X(t)$ for each t. (See [8].)

Let us define a *constructively Gaussian measure* as a probability measure μ on a commutative Polish group G, that can be represented as the distribution of some $X \in L_0(\gamma, G)$ satisfying (equivalent) conditions (a–c) of Lemma 5.3.

A constructively Gaussian measure is a Gaussian measure in the sense of Bernstein. (Indeed, 5.3(c) transfers Bernstein property of γ into Bernstein property of μ.)

A constructively Gaussian measure on a separable F-space is a shift of a symmetric Gaussian measure in the sense of Fernique. (Indeed, 5.3(c) and 5.4 transfer Fernique property of γ into Fernique property of μ shifted by $(-g)$.)

If a measure on a separable F-space is contained in a Gaussian semigroup, then it is constructively Gaussian (which is shown below by means of Theorem 1.8). The following fact is thus obtained.

5.5 Corollary. If a measure μ on a separable F-space is contained in a Gaussian semigroup, then some shift of μ is a symmetric Gaussian measure in the sense of Fernique, and μ is a Gaussian measure in the sense of Bernstein.

A Brownian motion in a separable F-space F determines a measure μ on the space $C_0\big([0,\infty), F\big)$ of all continuous functions $x : [0,\infty) \to F$ such that $x(0) = 0$. The space $C_0\big([0,\infty), F\big)$, equipped with the topology uniform on finite intervals, is also a separable F-space.

Corollary 1.10 *(final formulation).* For every Brownian motion in a separable F-space F, the corresponding measure μ on the space $C_0\big([0,\infty), F\big)$ is constructively Gaussian.

The corresponding Gaussian semigroup (μ_t) results from μ by applying evaluation maps $x(\cdot) \mapsto x(t)$. Each evaluation map is a continuous linear map $C_0\big([0,\infty), F\big) \to F$, therefore it sends a constructively Gaussian measure into another constructively Gaussian measure. Thus, 5.5 follows from 1.10.

PROOF OF COROLLARY 1.10. Let X be a Brownian motion in a separable F-space F. By Theorem 1.8, there exists a Brownian motion (X,Y) in $F \times \Delta$, where Δ is the Hilbert space, such that $\mathcal{F}_t^X = \mathcal{F}_t^Y$ for all t. Take a Borel function $f : C_0\big([0,\infty),\Delta\big) \to C_0\big([0,\infty),F\big)$ such that $f(Y) = X$ almost sure (here X is treated as a random variable $\Omega \to C_0\big([0,\infty),F\big)$).

Introduce two independent copies (X_1,Y_1) and (X_2,Y_2) of the Brownian motion (X,Y). It is easy to see that the process (X_1+X_2, Y_1+Y_2) is also a Brownian motion. Moreover, the process $t \mapsto \big(X_1(t/2)+X_2(t/2), Y_1(t/2)+Y_2(t/2)\big)$ is another copy of (X,Y), since $\big(X_1(t/2),X_2(t/2)\big)$ is distributed like $\big(X(t/2), X(t)-X(t/2)\big)$, and the same for (X,Y) pairs. For notational convenience, introduce $h : [0,\infty) \to [0,\infty)$ by $h(t) = t/2$; we see that $(X_1\circ h + X_2 \circ h, Y_1 \circ h + Y_2 \circ h)$ is distributed like (X,Y). Therefore, $f(Y_1\circ h+Y_2\circ h) = X_1\circ h+X_2\circ h$ almost sure. However, $f(Y_1) = X_1$ and $f(Y_2) = X_2$, thus, $f(Y_1\circ h + Y_2 \circ h) = f(Y_1)\circ h + f(Y_2)\circ h$. Define $g : C_0\big([0,\infty),\Delta\big) \to C_0\big([0,\infty),F\big)$ by $g(y) = f\big(2^{1/2} y \circ h\big) \circ h^{-1}$, then $g\big(2^{-1/2}(Y_1+Y_2)\big) = f(Y_1) + f(Y_2)$ almost sure. It means that f satisfies Condition (a) of Lemma 5.3. Though, the corresponding measure γ is not the standard Gaussian measure on $\mathbb{R}^\infty$, rather it is some Gaussian measure on $C_0\big([0,\infty),\Delta\big)$. However, it is well-known that each Gaussian measure on a *locally convex* F-space is linearly isomorphic mod 0 to the standard Gaussian measure on $\mathbb{R}^\infty$. Due to Lemma 5.3, f satisfies its Condition (c), therefore the distribution of $f(Y) = X$ is constructively Gaussian. ∎

5.6 NOTE. Linearity of F was not used (multiplications by $2^{\pm 1/2}$ were made in Δ, not in F). Thus, Corollary 1.10 (and its proof) remains true for all commutative Polish groups.

APPENDIX

PROOF OF LEMMA 3.1. First, let H be a Hilbert space of finite or countable dimension, X a Brownian motion in the unitary group $\mathrm{U}(H)$, and $(\mathcal{T}_t)$ the semigroup on V defined by (3.3). We extend $(\mathcal{T}_t)$ to the complexification of V, the space of all (not only Hermitian) trace-class operators, by complex linearity; still, (3.3) holds. Introduce a notation for one-dimensional operators on H: for any $h_1, h_2 \in H$

$$\langle\!\langle h_1, h_2 \rangle\!\rangle : H \to H\,, \qquad \langle\!\langle h_1, h_2 \rangle\!\rangle h = (h, h_2) h_1 \text{ for } h \in H\,.$$

(Physicists denote it by $|h_1\rangle\langle h_2|$.) Note some general rules:

$$\begin{aligned}
&\operatorname{tr}\big(\langle\!\langle h_1,h_2\rangle\!\rangle\big) = (h_1,h_2)\,,\\
&\big(\langle\!\langle h_1,h_2\rangle\!\rangle\big)^* = \langle\!\langle h_2,h_1\rangle\!\rangle\,,\\
&A\langle\!\langle h_1,h_2\rangle\!\rangle = \langle\!\langle Ah_1,h_2\rangle\!\rangle\,, \quad \langle\!\langle h_1,h_2\rangle\!\rangle A = \langle\!\langle h_1,A^*h_2\rangle\!\rangle \quad \text{for } A : H \to H\,,\\
&\langle\!\langle h_1,h_2\rangle\!\rangle\langle\!\langle h_3,h_4\rangle\!\rangle = (h_3,h_2)\langle\!\langle h_1,h_4\rangle\!\rangle\,,\\
&\operatorname{tr}\big(A\langle\!\langle h_1,h_2\rangle\!\rangle\big) = (Ah_1,h_2)\,.
\end{aligned}$$

Using the Hilbert-Schmidt scalar product $(B,C)_{\mathrm{HS}} = \operatorname{tr}(BC^*)$ for trace class operators $B,C : H \to H$, calculate a matrix element for $\mathcal{T}_t$: $\bigl(\mathcal{T}_t\langle\!\langle h_1,h_2\rangle\!\rangle, \langle\!\langle h_3,h_4\rangle\!\rangle\bigr)_{\mathrm{HS}} = \mathbb{E}\operatorname{tr}\bigl(X(t)\langle\!\langle h_1,h_2\rangle\!\rangle X^*(t)\langle\!\langle h_4,h_3\rangle\!\rangle\bigr) = \mathbb{E}\operatorname{tr}\bigl(\langle\!\langle X(t)h_1, X(t)h_2\rangle\!\rangle\langle\!\langle h_4,h_3\rangle\!\rangle\bigr) = \mathbb{E}\bigl(h_4, X(t)h_2\bigr)\bigl(X(t)h_1,h_3\bigr) = \mathbb{E}\bigl(X(t),\langle\!\langle h_3,h_1\rangle\!\rangle\bigr)_{\mathrm{HS}}\bigl(\langle\!\langle h_4,h_2\rangle\!\rangle, X(t)\bigr)_{\mathrm{HS}}$. In particular,

$$\bigl(\mathcal{T}_t\langle\!\langle h_1,h_2\rangle\!\rangle, \langle\!\langle h_1,h_2\rangle\!\rangle\bigr)_{\mathrm{HS}} = \mathbb{E}\bigl(X(t),\rho_1\bigr)_{\mathrm{HS}}\bigl(\rho_2, X(t)\bigr)_{\mathrm{HS}}\,, \tag{A.1}$$

where $\rho_k = \langle\!\langle h_k,h_k\rangle\!\rangle$.

From now on, H is assumed to be finite-dimensional. We have diffusion processes $A(t)$ in Δ and $X(t) = \exp(iA(t))$ in $\mathrm{U}(H)$; here Δ is the linear space of all Hermitian operators on H. We may write $A(t) = A_0(t) + \lambda(t)\cdot 1_H$, $\operatorname{tr} A_0(t) = 0$, $\lambda(t) \in \mathbb{R}$. For small t,

$$\begin{aligned}&X(t) = 1 + iA(t) - \tfrac12 A^2(t) + o(t)\,;\\ &\bigl(X(t),\rho\bigr)_{\mathrm{HS}} = 1 + i\bigl(A(t),\rho\bigr)_{\mathrm{HS}} - \tfrac12\bigl(A^2(t),\rho\bigr)_{\mathrm{HS}} + o(t)\end{aligned}$$

for $\rho = \langle\!\langle h,h\rangle\!\rangle$, $\|h\| = 1$. Therefore, for $\rho_k = \langle\!\langle h_k,h_k\rangle\!\rangle$, $\|h_k\| = 1$ $(k = 1,2)$,

$$\begin{aligned}&\mathbb{E}\bigl(X(t),\rho_1\bigr)_{\mathrm{HS}}\bigl(\rho_2,X(t)\bigr)_{\mathrm{HS}} =\\ &= 1 + i\bigl(\mathbb{E}\,A(t),\rho_1\bigr)_{\mathrm{HS}} - i\bigl(\rho_2,\mathbb{E}\,A(t)\bigr)_{\mathrm{HS}} +\\ &\quad + \mathbb{E}\bigl(A(t),\rho_1\bigr)_{\mathrm{HS}}\bigl(\rho_2,A(t)\bigr)_{\mathrm{HS}} - \tfrac12\bigl(\mathbb{E}\,A^2(t),\rho_1\bigr)_{\mathrm{HS}} - \tfrac12\bigl(\rho_2,\mathbb{E}\,A^2(t)\bigr)_{\mathrm{HS}} + o(t)\,.\end{aligned}$$

Separating real and imaginary terms (you see, $A(t)$ and ρ_k are Hermitian) and using (A.1) we conclude that the following two expressions are uniquely determined by the semigroup $(\mathcal{T}_t)$:

$$\frac{d}{dt}\Big|_{t=0}\bigl(\mathbb{E}\,A(t),\rho_1-\rho_2\bigr)_{\mathrm{HS}}\,; \tag{A.2}$$

$$\frac{d}{dt}\Big|_{t=0}\mathbb{E}\bigl(A(t),\rho_1\bigr)_{\mathrm{HS}}\bigl(A(t),\rho_2\bigr)_{\mathrm{HS}} - \frac12\frac{d}{dt}\Big|_{t=0}\bigl(\mathbb{E}\,A^2(t),\rho_1+\rho_2\bigr)_{\mathrm{HS}}\,. \tag{A.3}$$

Operators of the form $\rho_1 - \rho_2$ belong to the space Δ_0 of all traceless Hermitian operators, and Δ_0 is spanned by such operators. Thus, (A.2) means that the following is uniquely determined by $(\mathcal{T}_t)$:

$$\frac{d}{dt}\Big|_{t=0}\mathbb{E}\,A_0(t)\,. \tag{A.4}$$

Consider four expressions of the form (A.3): first, exactly as (A.3), that is, for the pair (ρ_1,ρ_2); second, the same for another pair, (ρ_3,ρ_4); also, for (ρ_1,ρ_4) and (ρ_3,ρ_2), the last two with the minus sign. Summing the four, we get

$$\mathbb{E}\bigl(A(t),\rho_1-\rho_3\bigr)_{\mathrm{HS}}\bigl(A(t),\rho_2-\rho_4\bigr)_{\mathrm{HS}}\,. \tag{A.5}$$

It means that the whole bilinear form describing the spread (at $t = 0+$) of $A_0(t)$ in Δ_0 is determined by $(\mathcal{T}_t)$. Together with (A.4) it proves that the generator of the Brownian motion $\left(\det X(t)\right)^{-1/d} X(t) = \exp\left(iA_0(t)\right)$ in $\mathrm{SU}(d)$ is uniquely determined by $(\mathcal{T}_t)$, which completes the proof of 3.1(a).

Turn to the semigroup (x_t), $x_t = \mathbb{E}\, X(t)$. We have $x_t = 1 + i\mathbb{E}\, A(t) - \frac{1}{2}\mathbb{E}\, A^2(t) + o(t)$; separating real and imaginary terms we conclude that the following two expressions are uniquely determined by (x_t):

$$\frac{d}{dt}\Big|_{t=0} \mathbb{E}\, A(t)\,, \tag{A.6}$$

$$\frac{d}{dt}\Big|_{t=0} \mathbb{E}\, A^2(t)\,. \tag{A.7}$$

However, $A(t) = A_0(t) + \lambda(t) \cdot 1_H$, and $\lambda(t) = (1/d)\mathrm{tr}\, A(t)$. Taking the trace of (A.6), we find $\frac{d}{dt}\big|_{t=0}\mathbb{E}\,\lambda(t)$. Further, $\mathbb{E}\, A^2(t) = \mathbb{E}\, A_0^2(t) + 2\mathbb{E}\,\lambda(t)A_0(t) + \mathbb{E}\,\lambda^2(t) \cdot 1_H$, and 3.1(a) ensures that $\mathbb{E}\, A_0^2(t)$ is determined by $(\mathcal{T}_t)$. Thus, the following is uniquely determined by (x_t) and $(\mathcal{T}_t)$:

$$2\frac{d}{dt}\Big|_{t=0} \mathbb{E}\,\lambda(t)A_0(t) + 1_H \cdot \frac{d}{dt}\Big|_{t=0} \mathbb{E}\,\lambda^2(t)\,. \tag{A.8}$$

The two terms can be separated by taking the trace. So, (x_t) and $(\mathcal{T}_t)$ determine

$$\frac{d}{dt}\Big|_{t=0} \mathbb{E}\,\lambda(t)\,, \quad \frac{d}{dt}\Big|_{t=0} \mathbb{E}\,\lambda^2(t)\,, \quad \frac{d}{dt}\Big|_{t=0} \mathbb{E}\,\lambda(t)A_0(t)\,; \tag{A.9}$$

these are all the infinitesimal characteristics of A (or X), besides characteristics of A_0 given by 3.1(a). ∎

References

1. S. Albeverio, V.N. Kolokol'tsov, O.G. Smolyanov, *Continuous quantum measurement: local and global approaches,* Reviews in Mathematical Physics **9** (1997), 907–920.
2. H. Araki and E.J. Woods, *Complete Boolean algebras of type I factors,* Publ. RIMS Kyoto Univ., Ser. A, **2** (1966), 157–242.
3. W. Arveson, *Continuous analogues of Fock space,* Memoirs AMS **80**:409 (1989).
4. W. Arveson, *E_0-semigroups in quantum field theory.* In: "Quantization, nonlinear partial differential equations, and operator algebra", Proc. Sympos. Pure Math. **59**, AMS 1996, pp. 1–26.
5. P. Baxendale, *Brownian motions in the diffeomorphism group I,* Compositio Mathematica **53** (1984), 19–50.
6. B.V.R. Bhat, *An index theory for quantum dynamical semigroups,* Trans. Amer. Math. Soc. **348** (1996), 561–583.
7. T.A. Brun, *Continuous measurements, quantum trajectories, and decoherent histories.* Report NSF-ITP-97-116, Inst. for Theor. Phys., Univ. of California, Santa Barbara, 1997. Electronic archive quant-ph/9710021. Submitted to Phys. Rev. A.
8. T. Byczkowski and A. Hulanicki, *Gaussian measure of normal subgroups,* Ann. Probab. **11** (1983), 685–691.

9. T. Byczkowski and T. Inglot, *Gaussian random series on metric vector spaces,* Math. Zeitschrift **196** (1987), 39–50.
10. Kai Lai Chung and J. B. Walsh, *Meyer's theorem on predictability,* Z. Wahrscheinlichkeitstheorie verw. Gebiete **29** (1974), 253–256.
11. E.B. Davies, *Quantum theory of open systems,* Academic Press, London, 1976.
12. J. Dixmier, *Les C^*-algèbres et leurs représentations,* Gauthier-Villars, Paris, 1969. English translation: *C^*-algebras,* revised edition, North-Holland 1982.
13. J. Feldman, *Decomposable processes and continuous products of probability spaces,* J. Funct. Anal. **8** (1971), 1–51.
14. P.R. Halmos, *Lectures on ergodic theory,* Chelsea, N.Y. 1956.
15. R.L. Hudson, *Quantum stochastic calculus, evolutions and flows.* In: "Quantization, nonlinear partial differential equations, and operator algebra", Proc. Sympos. Pure Math. **59**, AMS 1996, pp. 81–91.
16. K. Itô, *Brownian motions in a Lie group,* Proc. Japan Acad., **26** (1950), 4–10.
17. S. Itô, *Brownian motions in a topological group and in its covering group,* Rend. Circ. Mat. Palermo (2) **1** (1952), 40–48.
18. N.C. Jain, G. Kallianpur, *Norm convergent expansions for Gaussian processes in Banach spaces,* Proc. Amer. Math. Soc. **25** (1970), 890–895.
19. J. L. Kelley, *General topology,* Berlin, Springer 1955.
20. V.L. Klee, *Invariant metrics in groups (solution of a problem of Banach),* Proc. Amer. Math. Soc. **3** (1952), 484–487.
21. H. Kunita, *Stochastic flows and stochastic differential equations,* Cambridge Univ. Press, 1990.
22. G. Lindblad, *On the generators of quantum dynamical semigroups,* Commun. Math. Phys. **48** (1976), 119–130.
23. P. Malliavin, *Stochastic analysis,* Springer-Verlag, Berlin, 1997.
24. H. P. McKean, *Stochastic integrals,* Academic Press, New York 1969.
25. K.R. Parthasarathy, *An introduction to quantum stochastic calculus,* Birkhäuser, Basel, 1992.
26. K.R. Parthasarathy, *Quantum stochastic calculus,* Proc. Intern. Congress Math. 1994, pp. 1024–1035, Birkhäuser, Basel, 1995.
27. G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions,* Cambridge Univ. Press 1992.
28. M. Reed and B. Simon, *Methods of modern mathematical physics. I: Functional analysis.* Revised and enlarged edition, Academic Press, London 1980.
29. S. Rolewicz, *Metric linear spaces,* Warszawa 1972.
30. T. de la Rue, *Espaces de Lebesgue,* Lect. Notes Math. (Springer) **1557** (1993), 15–21.
31. H. Satô, *Souslin support and Fourier expansion of a Gaussian Radon measure,* Lect. Notes Math. (Springer) **860** (1981), 299–313.
32. A. Shnirelman, *On the non-uniqueness of weak solution of the Euler equation.* Preprint IHES/M/96/31.
33. A. V. Skorokhod, *Asymptotic methods in the theory of stochastic differential equations,* AMS, Providence, R.I., 1989 (translated from Russian).

34. B.S. Tsirelson and A.V. Vershik, *Examples of nonlinear continuous tensor products of measure spaces and non-Fock factorizations,* Reviews in Mathematical Physics **10** (1998), 81–145.
35. K. Yosida, *On Brownian motion in a homogeneous Riemannian space.* Pacific J. Math. **2** (1952), 263–270.