AT 83 00043

UWThPh-1982-5

DYNAMICAL AND HAMILTONIAN DILATIONS OF STOCHASTIC PROCESSES

B. Baumgartner and H.-R. Grümm (67110004T)

Institut für Theoretische Physik

Universität Wien

Abstract

This is a study of the problem, which stochastic processes could arise from dynamical systems by loss of information. The notions of "dilation" and "approximate dilation" of a stochastic process are introduced to give exact definitions of this particular relationship. It is shown that every generalized stochastic process is approximately dilatable by a sequence of dynamical systems, but for stochastic processes in full generality one needs nets.

1. Introduction

To study the interrelations between deterministic and stochastic quantum evolutions, there has been done much work on the questions of unitary dilations of quantum dynamical semigroups [1]. The analogous problem for classical evolutions can be treated with the well-known Kolmogorov-construction which enables one to construct to any Markov process with an invariant measure a one parameter group of measure preserving transformations of path-space [2]. In this paper, we study the classical case, but our investigation differs from the mentioned one in two points: On one hand, we do not limit ourselves to Markov processes. On the other hand, our prototype for a deterministic evolution will be an ordinary differential equation on some differential manifold (= "phase space"). We try, therefore, to dilate a stochastic process to a dynamical process, by which we understand a dynamical system (= a deterministic evolution as above) together with a distinct observable and a distinct state on the space of observables.

Definition 1. A dynamical process P is a quadruple (M, S_t, X, ρ) where

i) M is a differentiable manifold;

ii) S_t is a differentiable flow on M

iii) X is a real C^1-function on M

iv) ρ is a Borel measure on M.

There are obvious multi-dimensional generalizations of this definition which, however, do not lead to a qualitatively different situation. Therefore we restrict ourselves to just one observable.

Definition 2 A Hamiltonian process is a quadruple $(T^*(M), S_t, X, \rho)$ where

i) $T^*(M)$ (the phasical phase space) is the cotangent bundle of a manifold M. More generally, M could be any symplectic manifold. For these notions and their physical significance, see [3] or [4].

ii) S_t is a Hamiltonian[1)] flow on $T^*(M)$;

iii), iv) as above.

1) Somewhat loosely, we call a flow Hamiltonian if it consists of canonical transformations, i.e. those respecting the natural symplectic structure on $T^*(M)$. It is well known that such flows are locally generated by functions in the sense of being defined by Hamilton's equations. All our manifolds will be simply connected, thus there is no difference between "local" and "global".

We consider Hamiltonian processes because the classical time evolution of a physical system is often required to be of Hamiltonian character. The condition of differentiability of X (the observable) will be relaxed only in the second part of chap. 3 to show that we would obtain some counterintuitive results without it.

As is well known, a stochastic process can be defined by the hierarchy of its correlation measures $\mu(X_{t_1} \in A_1; \dots X_{t_n} \in A_n)$ where the A_i are Borel subsets of $\mathbb{R}$ or equivalently by its expectation functionals $E(f(X_{t_1}; \dots X_{t_n})$ where f is a bounded continuous function on $\mathbb{R}^n$. We will also use the characterization of stochastic processes by their characteristic functions which are the Fourier-Stieltjes transforms of their correlation measures:

$$\chi(t_1 \dots t_n, \lambda_1 \dots \lambda_n) = E(\exp i \sum_{k=1}^{n} \lambda_k x_k) . \tag{1.1}$$

We observe that a dynamical process yields immediately a stochastic process by setting $X_{t_i} = X \circ S_{t_i}$ ("observation of the single observable X as time goes by"). If this leads to the same correlation measures/ expectation functionals/characteristic functions as those of a <u>given</u> stochastic process, i.e.

$$\rho(X \circ S_{t_1} \in A_1; \dots X \circ S_{t_n} \in A_n) = \mu(X_{t_1} \in A_1, \dots X_{t_n} \in A_n) \ \forall\, n, \{A_k\}, \{t_k\}^{2)} \tag{1.2}$$

$$\rho(f(X \circ S_{t_1}, \dots X \circ S_{t_n})) = E(f(X_{t_1}, \dots X_{t_n})) \ \forall\, n, \{t_k\}, f \in C_B(\mathbb{R}^n) \tag{1.3}$$

$$\rho(\exp(i \sum \lambda_k (X \circ S_{t_k})) = \chi(t_1, \dots t_n, \lambda_1, \dots \lambda_n) \ \forall\, n, \{t_k\}, \{\lambda_k\} \tag{1.4}$$

we call this a dilation.

<u>Definition 3</u>. A dynamical Hamiltonian process is a dilation of a given stochastic process if one of the equivalent conditions (1.2), (1.3) or (1.4) holds.

---

2) All symbols on the r.h.s. of equations (1.2) to (1.5) refer to the given stochastic process; those on the l.h.s. to the dynamical process.

We will also consider dilations of generalized stochastic processes (GSP). Loosely speaking, a GSP differs from an ordinary stochastic process only in that the correlation measures are only defined when "smeared out" in time with test-functions belonging to some test-function space $\mathcal{E}$. It is therefore characterized by its characteristic functional $\chi[\phi] = E(\exp i\phi)$, $\phi \in \mathcal{E}$ (see Minlos [5] or Gelfand IV [6] for details).

Definition 4. A dynamical or Hamiltonian process is a dilation of a given GSP indexed by $\mathcal{E}$, if

$$\rho(\exp i \int \phi(t)\, X \circ S_t\, dt) = \chi[\phi] \, . \tag{1.5}$$

Finally, we introduce the dilation of one dynamical process by another one:

Definition 5. A process $\tilde{P} = (\tilde{M}, \tilde{S}_t, \tilde{X}, \tilde{\rho})$ is a dilation[3] of the process $P = (M, S_t, X, \rho)$ if there exists a submersion π from $\tilde{M}$ onto M such that

i) $$\begin{array}{ccc} \tilde{M} & \xrightarrow{\tilde{S}_t} & \tilde{M} \\ \pi \downarrow & & \downarrow \pi \\ M & \xrightarrow[S_t]{} & M \end{array}$$ is commutative for all t

ii) $\tilde{X} = X \circ \pi$

iii) $\rho = \pi^*(\rho)$ (π^* is the map on measures induced by ρ).

Intuitively speaking, P arises from $\tilde{P}$ by projecting from M onto some lower-dimensional state space M, thereby concentrating only on those variables which could serve as coordinates on M.

3) P is also called a factor of $\tilde{P}$ in the literature.

2. Dilations of Processes Indexed by a Finite Set

A process indexed by a finite set of times $t_1 < t_2 < \ldots < t_N$ can be defined by its correlation measure on $\mathbb{R}^N$, whose coordinates are to be identified with the X_{t_i}. Suppose first that the t_i are equally spaced, i.e. $t_i - t_{i-1} = \tau$. The permutation P of the basis vectors $e_{i+1} \to e_i, e_1 \to (-)^N e_N$ (or $X_{t_n} \to X_{t_{n+1}}$, $X_{t_N} \to (-)^N X_{t_1}$) defines a special orthogonal transformation of $\mathbb{R}^N$ which, by the connectivity and compactness of SO(N) can be embedded in an orthogonal flow (= one-parameter subgroup of SO(N)) S_t such that $S_\tau = P$. The dilated dynamical system is then $(\mathbb{R}^N, S_t, X_1, \mu)$. If the t_i are not equally spaced, we just introduce a different time scale. The flow S_t can be decomposed as a direct sum of two-dimensional rotations, the periods of which are certain multiples of τ/N. There are therefore choices of e.g. $\tau = \max(t_i - t_{i-1})$ which guarantee the linear independence of the set $\{S_{t_i} e_1\}$. Let A be the linear transformation taking $S_{t_i} e_1$ into e_i. The dynamical system dilating the given process can be chosen as $(\mathbb{R}^N, S_t, x_1, A^*\mu)$. Since a two-dimensional rotation is a Hamiltonian flow (physicists should think of the harmonic oscillator), we already have a Hamiltonian dilation for even N (for odd N add one extra dimension).

Theorem 2.1. Any process indexed by a finite set of times $\{t_1 \ldots t_N\}$ can be dilated to a (C^∞) dynamical process in N dimensions and to a Hamiltonian process with N/2 resp. (N+1)/2 "coordinates" and the same number of "momenta".

In the construction above we have tried to be economical with respect to the number of dimensions. An alternative construction could proceed as follows: Any discrete (equal-times) dynamical system can be embedded into a flow in one extra dimension by suspension (see f.i. Arnold-Avez [7]). A change of time-scale is obtained by treating time as an additional coordinate (again one extra dimension); if the required change is given by a strictly increasing C^1-function h on $\mathbb{R}$ ($t = h(s)$ and the original flow by $dx/ds = V(x)$, the new flow is just $ds/dt = 1/h'(s)$, $dx/dt = V(x)/h'(s)$. Finally, any flow can be dilated to a Hamiltonian one by doubling the number of dimensions: this step is known in classical mechanics under the name of "point transformations as a special case of canonical transformations". If S_t is any flow on M, the induced flow T^*S_t on the "phase space" $T^*(M)$ is automatically Hamiltonian. The required projection is the natural projection of the cotangent bundle onto its base space.

3. The Minimal Number of Dimensions

For a generic stochastic process indexed by a finite set the measure on path space (which is simply $\mathbb{R}^N$ with coordinates X_{t_i}) is not supported by a submanifold of $\mathbb{R}^N$ of lower dimension. Suppose that M is the m-dimensional phase space of a dynamical process P which is a dilation of the given process and that the C^1-observable defining the dilation is X. The map $M \to \mathbb{R}^N$, $m \to (X \circ S_{t_1}(m), \ldots X \circ S_{t_N}(m))$ is C^1; thus by Sard's theorem (in the strengthened version; see Abraham-Robbin [8]) its image cannot have dimension higher than m. Since obviously the support of μ has to be contained in this image, we conclude that, generically, the dimension of M cannot be chosen smaller than N.

Theorem 3.1. **A dynamical process $P = (M, S_t, X, \rho)$ which is a dilation of a given stochastic process with an associated measure on path space μ, cannot have lower dimension than the smallest dimension of a submanifold containing the support of μ.**

This holds a fortiori if the defining observable is C^r or C^∞, but the C^1-requirement cannot be disregarded. Allowing a merely continuous or measurable observable X, the situation changes drastically. The following construction shows that every stochastic process P indexed by Z can be dilated to a flow on $\mathbb{R}$ using a Borel measurable X. If the process is almost surely locally bounded, i.e. the measure on path space is supported by a hypercube of the form $|X_t| < c_t < \infty$, X may even be chosen as a continuous function.

This goal is achieved with the help of a continuous mapping $\Phi = (\Phi_k)$ from the unit interval $[0,1]$ onto path space $S = \dot{\mathbb{R}}^Z$ (or $\times_{t \in Z} [-c_t, c_t]$);

we take a compactification of $\mathbb{R}$ for technical reasons. Examples of such mappings are described in the Appendix. As the next step we show the existence of a measure ρ on S which maps under X (in the sense of the map induced by Φ on measures) to the normed measure on the cylinder sets of path space μ associated to P. ($\rho(\Phi^{-1}(B)) = \mu(B)$ for any cylinder set B.) This is guaranteed by the Hahn-Banach theorem, once we have identified measures with linear functionals on suitable function spaces. The map Φ induces linear transformations Φ^*: $B(S) \to B([0,1])$ resp. $C(S) \to C([0,1])$, the Borel measurable and bounded resp. continuous functions on S or $[0,1]$. We can define ρ first on $\Phi^* B([0,1])$ resp. $\Phi^* C([0,1])$: Since Φ is onto, Φ^* is injective; it is also continuous in the supremum norm, as well as its left-inverse Φ^{*-1}; thus we can put $\rho = \mu \circ \Phi^{*-1}$, a continuous positive functional on $\Phi^* B([0,1])$ resp. $\Phi^* C([0,1])$. This functional can be extended to the whole algebra by Hahn-Banach.

To finish the construction, we take $\mathbb{R}$ as the phase space of the dilating dynamical process and we define the flow by $S_t(r) = t + r$. The observable X is given by the components of Φ and we just showed the existence of a suitable ρ. We write

$$X(k + \lambda) = \Phi_k(\lambda) \; ; \quad k \in \mathbb{Z}, \quad \lambda \in [0,1) \; .$$

This implies $S_k^* X = X \circ S_k = \Phi_k$ on the unit interval and we can check the equality of the expectation functionals ($f \in C(\mathbb{R}^N)$):

$$\mu(f(X_{t_1}, \dots X_{t_N})) = \rho(f \circ \Phi) = (f(X \circ S_{t_1}, \dots X \circ S_{t_n})) \; .$$

By the construction at the end of chap. 2 we even achieve a Hamiltonian dilation; taking the measure $\rho \times \delta(p-1)$ on $T^*(\mathbb{R}) = \mathbb{R}^2$ the process turns out to be the well known "free particle in one dimension".

4. Approximate Dilations of Processes

For the generic stochastic process, "chance intervenes at every instant of time" [9], so that the relevant path space will be infinite-dimensional, i.e. it cannot be finitely parametrized. A dynamical or Hamiltonian system with a finite number of degrees of freedom cannot serve as a dilation of such a process, unless one dismisses the algebra of C^{L}observables and uses the rather unphysical construction of the last section. Keeping to differentiable systems and observables, one has to be content with approximate dilations. From the physical point of view, the natural topology on any set of stochastic processes is the weak topology since any finite measurement can only determine a finite set of correlation functions. We therefore should check the convergence of correlation measures or of characteristic functions (individually!). In the following we call a process dilatable if it has a finite dimensional Hamiltonian system as a dilation.

Theorem 4.1. Every stochastic process P is the weak limit of a net of dilatable processes, i.e. the dilatable processes are weakly dense.

Proof. Consider any finite set of times $\{t_1,t_2,\ldots t_N\}$. According to section 2 we can choose a Hamiltonian process $S_{t_1,t_2,\ldots t_N}$ and a measure $\rho_{t_1,t_2,\ldots t_N}$ on its phase space dilating the restriction of P to the finite sets of times, i.e. yielding the same characteristic function $\chi_{t_1,t_2,\ldots t_N}$. We partially order the obtained processes by inclusion of their finite sets of times; this defines a net of processes converging to P.

We had to formulate this result in terms of net convergence because of the non-metrizability of the space of all stochastic processes. In contrast, the limit procedures usually encountered in physics do not use general nets but sequences (= nets with a denumerable basis), e.g. the thermodynamic limit, scaling limits etc. This suggests that the space of all stochastic processes is far too big for ordinary applications and the next results confirms this.

Theorem 4.2. There exist stochastic processes which are not limits of sequences of dilatable processes.

Proof. We give a simple counterexample. Let P be a stochastic process whose measure in path space is concentrated on a single path $y(t)$; let y be a function not in the first Baire category. Take $f(x) = \arctan(x)$ (or any homeomorphism of $\mathbb{R}$ on an open bounded interval), then $E(f(x_t)) = f(y(t))$. This cannot be the pointwise limit of a sequence $\rho_n(f(x_t))$ coming from Hamiltonian (or even dynamical) systems because these functions of t are all continuous.

To realize the pathological nature of this counterexample, let us remind ourselves that a function of, e.g. second Baire class looks like Dirichlet's function which is $= 0$ on the irrationals and $= 1$ on the rationals. Smeared with test functions, it is indistinguishable from the zero function. We therefore turn to generalized stochastic processes in the hope for better results.

Theorem 4.3. Every generalized stochastic process P indexed by S is a weak limit of a sequence of (GSPs induced by) dilatable processes.

Note. The relevant property of S is the following: The L^2-inner product $\langle\ |\ \rangle_2$ is continuous and there exists a sequence $\{h_n\}$ in S, orthonormal in L^2, such that for all $f \in S$,

$$f = \sum_n \langle h_n | f\rangle_2 \, h_n \ , \qquad (*)$$

the sum converging in the topology of S. (Take $e^{-x^2/2}$ times the Hermite polynomials.) The theorem and proof below hold for any nuclear test function space with corresponding sequence $\{h_n'\}$.

Proof. We define the approximating GSP's by their characteristic functionals $\chi_N(f) = \chi(P_N f)$ where χ denotes the characteristic functional of P and $P_N f$ is the N-th partial sum in (*), i.e. the P_N form a sequence of finite-rank projectors pointwise converging to the identity in S. To make sure that χ_N is indeed a characteristic functional, one checks continuity, positive definiteness and $\chi_N(0) = 1$; this is almost trivial because of the corresponding properties of χ. The weak convergence of GSP's is equivalent to $\chi_N(f) \to \chi(f)$ for each $f \in S$. It remains to find a dynamical system which is a dilation of the process defined by χ_N: We take $M = \mathbb{R}^{2N}$ coordinatized by (x_n, y_n), S_t to be the flow S_t: $x_n \to x_n + t$,

$y_n \to y_n$, and $X = \sum_n y_n h(x_n)$. S_t is already Hamiltonian w.r.t. the natural symplectic structure on $\mathbb{R}^{2N}$. We identify the dual space of P_N first via the L^2-scalar product with $\{\sum_{n=1}^{N} y_n h_n(t)\}$ and further with $\mathbb{R}^N$ (coordinates y_n). χ_N induces a measure $\tilde{\rho}$ on this dual space and therefore on $\mathbb{R}^N$, too, by the formula

$$\int \exp(i \sum_{n=1}^{N} y_n \langle h_n | f \rangle_2) d\tilde{\rho}(y) = \chi_N(f) \ . \qquad (4.1)$$

We multiply $\tilde{\rho}$ by the Dirac measure at the origin of the x-coordinates to obtain the required state for our dynamical system.

Appendix I

There exists a "Peano curve", $Y(s)$, a map from the unit interval to the Hilbert cube $\prod_{n=1}^{\infty} [0,1] = \{y_1, y_2, \ldots, 0 \leq y_n \leq 1\}$ which is continuous and surjective. We give a concrete example.

The curve is best defined in geometrical terms by a sequence of approximants. The first approximation is constant. In the n-th step, the range of y_m for $1 < m < n$ is divided into 2^{n-m} intervals of equal length, so that the Hilbert cube is divided into $2^{n(n-1)/2}$ hyperboxes B_k^n. The n-th approximation to the Peano curve is the polygonal path through the centers of all B_k^n. Going from n to n+1, each hyperbox B_k^n is divided into 2^n hyperboxes B_ℓ^{n+1} $(2^n(k-1) < \ell \leq 2^n k)$. These are ordered in an appropriate way: the first one touches $B_{2^n(k-1)}^{n+1}$ (assumed to be defined already) or contains $(0,0,0\ldots)$ if $k = 1$, the last one lies between $B_{2^n(k-1)+1}^{n+1}$ and B_{k+1}^n or contains $(1,0,0\ldots)$ if k is maximal. It is easily seen by induction that the centers of the B^{n+1} can be joined by a polygonal path in B_k^n (apply a succession of "up, over and down").

The parametrization is chosen in a way such that the approximant $Y^{(n)}(s)$ stays inside B_k^n if $s \in [2^{-n(n-1)/2}(k-1), 2^{-n(n-1)/2}k] = I_{n,k}$. Then $Y^{(n)}$ converges to an Y, since the relevant boxes form a nested sequence. Each component $y_m(s)$ is continuous which yields the continuity of Y: for $|s-s'| \leq 2^{-(n-1)n/2}$ we have $|y_m(s)-y_m(s')| \leq 2^{m-n}$. That Y is onto follows easily: Each point y in the Hilbert cube is the intersection of a nested sequence of boxes B_k^n. The corresponding $I_{n,k}$ then intersect in a single point s which is mapped to y by Y.

The map X from the unit interval to path space is constructed from Y with the help of some reparametrization. One defines $n(\ell) = 2|\ell| + \theta(\ell)$, $n(0) = 1$, mapping the integers to the natural numbers and chooses a positive δ less than 1. If the paths are locally bounded we can take

$$x_\ell(\delta s) = c_\ell\, y_{n(\ell)}(s) \qquad (c_\ell \text{ appropriate bounds})$$

else

$$x_\ell(\delta s) = \begin{cases} \arctan\left(\pi(y_{n(\ell)}(s) - 1/2)\right) \\ 0 \text{ if } y_{n(\ell)}(s) = 0 \text{ or } 1 \end{cases} .$$

In both cases, $x_\ell(1) = 0$ and for $s \in (\delta, 1)$, $x_\ell(s)$ interpolates linearly between $x_\ell(\delta)$ and 0.

Appendix II

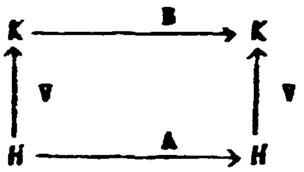
The motivation for calling the situation of Def. 5 - especially diagram i) - a dilation can be explained as follows:

As a first step, we apply the contravariant functor $X \to C_0(X)$ (X a locally compact space, $C_0(X)$ the C^*-algebra of continuous functions on it vanishing at infinity) to the diagram. We obtain

$$\begin{array}{ccc} C_0(\tilde{M}) & \xrightarrow{\tilde{S}_t^*} & C_0(\tilde{M}) \\ \uparrow \pi^* & & \uparrow \pi^* \\ C_0(M) & \xrightarrow{S_t^*} & C_0(M) \end{array} \qquad \text{(II.1)}$$

where $\tilde{S}_t^*$, S_t^* and π^* denote the naturally induced maps. π^* are norm-preserving embeddings. This reformulation permits us to regard Markovian stochastic processes on the same footing by allowing $\{S_t^*\}$ to be a semigroup instead of a group.

Secondly, we recall the definition of the dilation of an operator (Halmos [10]) which can be recast in diagramatic language: B is a dilation of A if the following diagram commutes:

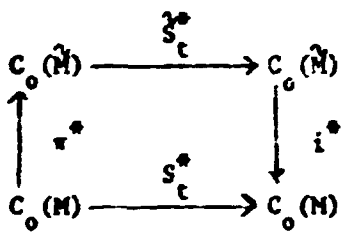


(II.2)

K, H Hilbert spaces, V an isometry. The equivalence follows by considering a left-inverse V^* of V: $A = V^*BV$.

If the dilation is a unitary power dilation - i.e. we can write the group of unitaries $\{U^n\}$ above and the semigroup $\{A^n\}$ below - the analogy to the first diagram is striking; we recall that there exists a "continuous diagram II.2" as well with continuous one-parameter groups instead of discrete ones.

It can be even strengthened if there is a natural right-inverse i of π, i.e. an embedding of M in $\tilde{M}$. This holds for the cotangent bundle (see chap. 2) by embedding M as the zero-section of $T^*(M)$. We then can rewrite (II.1) as



(II.3)

or $S_t^* = i^* \tilde{S}_t^* \pi^*$. This i^* is even a conditional expectation in the sense of [1].

These considerations suggest the following "dictionary" for the special case of Markov processes:

one-parameter groups	⟷	classical dynamical systems
semi-groups	⟷	classical Markovian stochastic processes
unitary one-parameter groups	⟷	classical Hamiltonian processes

For the last correspondence we need only mention the keywords "Liouville's theorem" and "conservation of probability".

References

[1] Evans, D.E., and Lewis, J.T.: Dilations of irreversible evolutions in algebraic quantum theory. Commun. Dublin Institute for Advanced Studies, A 24, 1977

[2] Feller, W.: An Introduction to Probability Theory and Its Apllications, vol. II. New York-London-Sidney: Wiley & Sons, 1966

[3] Thirring, W.: Lehrbuch der mathematischen Physik 4. Springer-Verlag, Wien-New York, 1980

[4] Abraham, R.: Foundation of Mechanics, Benjamin, New York, 1967

[5] Minlos, R.A.: Generalized random processes and their extension to a measure. Trudy Moscov. Mat. Obsc 8, 497-518 (1959). Translated in: Selected Translations in Mathematical Statistics and Probability 3, 291-313 (1962)

[6] Gelfand, I.M., Vilenkin, N.Ya.: Generalized Functions, vol. 4. New York: Academic Press 1964

[7] Arnold, V.I., and Avez, A.: Ergodic Problems of Classical Mechanics, Benjamin, New York 1968

[8] Abraham, R., and Robbin, J.: Transversal Mappings and Flows, Benjamin, New York 1967

[9] Lévy, P.: Processus stochastiques et mouvement brownien. Paris: Gauthier-Villars 1948

[10] Halmos, P.R.: A Hilbert Space Problem Book. New York: Springer 1974