Український математичний вісник
Том 1 **(2004)**, № 4, 449 – 484

Nonlinear Calculus of Variations for Differential Flows on Manifolds: Geometrically Correct Introduction of Covariant and Stochastic Variations

Alexander Val. Antoniouk and Alexandra Vict. Antoniouk

(*Presented by S. Ya. Makhno*)

Abstract. We consider flows, generated by nonlinear differential equations on manifold that could also contain random terms and correspond to the second order parabolic equations. We demonstrate that the rigorous statement of the regularity problems for differential flows on non-compact manifolds requires the geometrically rigorous revision of definition of the high order variation with respect to the initial data and parameters.

The main attention is devoted to the study of influence of the geometry and nonlinearities of coefficients on the regularity properties. To reach this aim we use the nonlinear symmetries of high order differential calculus and study a set of corresponding nonlinear estimates on variations.

The arising conditions on regularity generalize the Krylov-Rosovskii-Pardoux conditions from linear space to the manifold setting. They also lead to the smooth and smoothing properties of associated Feller semigroups.

2000 MSC. 35K05, 47J20, 53B21, 58J35, 60H07, 60H10, 60H30.

Key words and phrases. Heat parabolic equations, nonlinear diffusions, nonlinear estimates on ordinary and stochastic variations, raise of regularity, non-singular representations of semigroup derivatives.

1. Introduction into geometrically correct generalization of high order variations of differential flows on manifolds.

Though there already exist constructions that agree the geometry with the second order differentials, generated by Itô formula [10, 12, 15],

Received 14.04.2004
The research was supported by A.von Humboldt Foundation (Germany)

ISSN 1810 – 3200.

see also e.g. [13, 18, 23] and discussions in [4, 6], the aim to establish the principal relations between advanced objects of differential geometry (like curvature) and related regular properties of general differential flows on manifold is not achieved yet. The question how curvature should arise in the regularity properties is already asked in literature [13, 21]. In known approaches the corresponding high order variations are introduced by writing the covariant and stochastic derivatives of differential flows, see e. g. [10, 12, 18] and references therein. How we will soon show, such variations of nonlinear flows are actually non-invariant under coordinate transformations in vicinities, in which travels process y_t^x, i. e. do not form invariant geometric objects.

The question still remains: what relation between the coefficients of differential equation and geometry should be imposed to lead to any order regularity of processes on manifolds. We are going to demonstrate that the knowledge of simple symmetries of variations and a little work to make the geometrically correct definition of the high order ordinary and stochastic variations leads to the final answers.

We aim this paper to develop results of [2, 3] from the linear base space to the manifold setting. The main problem concerns the study of regular properties in the general nonlinear case, i. e. when the classical Cauchy-Liouville-Picard scheme, primarily developed for the Lipschitz or quasi-linear equations, does not work.

We consider differential equation on the oriented smooth connected Riemannian manifold M without boundary, that could also contain random terms

$$y_t^x = x + \int_0^t A_0(y_s^x)\, ds + \sum_\sigma \int_0^t A_\sigma(y_t^x)\, \delta W_t^\sigma. \tag{1.1}$$

Let us specially mark that we adopt the inconvenient for stochastic theory notation y_t^x for the solution of this differential equation. This is because the results of article still hold for ordinary differential equations on manifolds, and, therefore, do not follow the traditional arguments, aimed to create the stochastic differential geometry [10, 12, 13].

Above A_0, A_α represent smooth globally defined vector fields, W_t^σ denotes a family of one dimensional independent Wiener processes, δW^σ means Stratonovich differential, range of index σ corresponds to the dimension of manifold. Equation (1.1) is understood in a sense, that for any C^3 function on manifold the following equation

$$f(y_t^x) = f(x) + \int_0^t (A_0 f)(y_s^x)\, ds + \sum_\sigma \int_0^t (A_\sigma f)(y_t^x)\, \delta W_t^\sigma$$

holds as stochastic equation in $I\!R^1$. In particular, one can take $f(x) = x^k$ to find its local coordinates representations.

The corresponding semigroup

$$(e^{t\mathcal{L}} f)(x) = \mathbf{E}\, f(y_t^x) \tag{1.2}$$

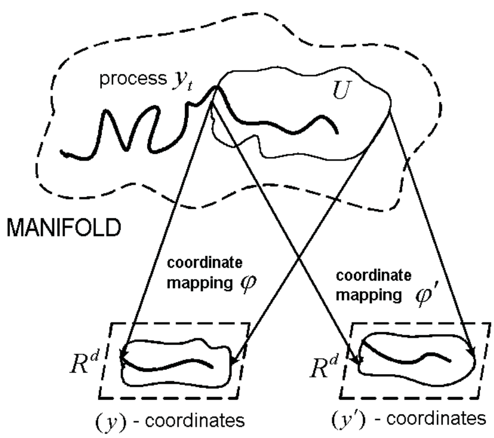
provides the solutions to the parabolic Cauchy problem

$$\frac{\partial}{\partial t} u(t,x) = \mathcal{L} u(t,x), \quad \mathcal{L} = \frac{1}{2} \sum_\sigma \nabla^2_{A_\sigma} + \nabla_{A_0}, \tag{1.3}$$

and gives therefore a set of actual applications to the problems of infinite dimensional functional and nonlinear analysis, stochastics, mathematical physics, differential geometry and operator theory. Henceforth we use notation $\nabla_A = \langle A, \nabla \rangle$ for covariant derivatives in direction of vector field A. We also remark that a simpler non-stochastic case of ordinary differential equation arises by taking $A_\sigma = 0$ and omitting the symbol of expectation $\mathbf{E}$ on Wiener measure. However the stochastic derivatives and raise of smoothness properties arise only in the stochastic case, i. e. only for the second order differential operators.

Concerning the problem of differentiability of equation (1.1) with respect to the initial data or random parameter (on Wiener space, on which process y_t^x depends), it is already clear how to write the first order variations. But how one should introduce the high order variations?

Suppose that some process on manifold y_t (of diffusion or any other nature) travels over manifold and enters some vicinity $U \subset M$ with coordinate functions $\varphi = (\varphi^i)_{i=1}^{\dim M}$, $\varphi : U \to I\!R^{dimM}$. Then one can speak about the coordinates of process $y_t^i = (\varphi^i) \circ y_t$ when it stays in this vicinity.



Let $\mathcal{D}$ be some first order differentiation operation, correctly defined on process y_t. It could be of any nature, like partial derivative ∂_x or stochastic derivative with respect to the random parameter, the principal moment is that the first order differentiation obeys chain rule

$$\mathcal{D}(f \circ y) = (f' \circ y)\, \mathcal{D}y.$$

Because the local coordinate changes $y^{i'} = \varphi^{i'}(y_t) = (\varphi^{i'} \circ \varphi^{inv})(y^i)_{i=1}^{dimM}$ represent a particular case of locally defined functions, one gets rule

$$\mathcal{D}y^{m'} = \frac{\partial y^{m'}}{\partial y^m} \mathcal{D}y^m.$$

Therefore

> the expression $\mathcal{D}y$ becomes a **vector field** with respect to the "coordinate" changes $(y) \to (y')$ of process "variable" y,

though, of course, the process y_t does not determine some coordinate system, like local coordinate mappings φ, φ' do.

By similar to the classical differential geometry arguments, related with the construction of covariant derivatives,

> there is no other way to define correctly the high order derivatives $(\widetilde{\mathcal{D}})^i y$, but consider additional terms with connection $\Gamma(y_t)$ depending on flow y_t^x.

Therefore the correct recurrent definition of the invariant high order derivative $(\widetilde{\mathcal{D}})^i$ becomes

$$\widetilde{\mathcal{D}}y^m = \mathcal{D}y^m, \quad \widetilde{\mathcal{D}}[\,(\widetilde{\mathcal{D}})^i y^m\,] = \mathcal{D}[\,(\widetilde{\mathcal{D}})^i y\,] + \Gamma_{p\ q}^{\ m}(y)\,[\,(\widetilde{\mathcal{D}})^i y^p\,]\mathcal{D}y^q, \qquad (1.4)$$

here also arise factors $\mathcal{D}y^q$ in the term with $\Gamma(y_t)$.

These additional terms in definition of higher order derivatives $\widetilde{\mathcal{D}}^n$ guarantee the preservance of vector transformation law with respect to the $(y) \to (y')$ coordinate transformations:

$$(\widetilde{\mathcal{D}})^n y^{m'} = \frac{\partial y^{m'}}{\partial y^m}\,(\widetilde{\mathcal{D}})^n y^m, \quad \forall n \geq 1.$$

Such $(y) \to (y')$ invariance, or, if one returns to the very beginning, the invariance with respect to changes of local coordinates $(y) \to (y')$ in vicinity, where process y_t stays, imposes a new purely geometric requirement. This requirement is very important, because it permits to introduce the invariant norms $|\widetilde{\mathcal{D}}y_t^x|_{T_{y_t^x}M}$ by traces with the metric tensor $g_{ij}(y_t^x)$ of image coordinate (y_t^x). After that the question of a priori estimates on the regularity of solutions becomes **well-posed**.

Consider, for example, the correct construction of high order variations of process y_t^x (1.2) with respect to the initial data. One should first note that first order variation $\dfrac{\partial (y_t^x)^m}{\partial x^k}$ represents a vector field on index m for $(y) \to (y')$ "coordinate" transformations and covector field on index k for $(x) \to (x')$ coordinate changes.

From arguments above it follows that the definition of *geometrically invariant high order variations* with respect to the initial data x must include terms with $\Gamma(x)$ and $\Gamma(y)$ to guarantee the preservance of tensorial character on both image $(y) \to (y')$ and domain $(x) \to (x')$ coordinate changes of mapping $x \to y_t^x$. Recurrently the high order variation $\mathbb{V}_\gamma^x y^m = \mathbb{V}_{j_n}^x \dots \mathbb{V}_{j_1}^x y^m$, $\gamma = (j_1, \dots, j_n)$, is defined from the first variation by

$$\mathbb{V}_k^x y^m = \frac{\partial (y_t^x)^m}{\partial x^k},$$

$$\begin{aligned} \mathbb{V}_k^x(\mathbb{V}_\gamma^x y^m) &= \nabla_k^x(\mathbb{V}_\gamma^x y^m) + \Gamma_{p\ q}^{\ m}(y_t^x)\mathbb{V}_\gamma^x y^p \frac{\partial y^q}{\partial x^k} \\ &= \underbrace{\partial_k^x(\mathbb{V}_\gamma^x y^m) - \sum_{j\in\gamma} \Gamma_{k\ j}^{\ h}(x)\mathbb{V}_{\gamma|_{j=h}}^x y^m}_{\textbf{old covariant derivative}} + \underbrace{\Gamma_{p\ q}^{\ m}(y_t^x)\mathbb{V}_\gamma^x y^p \frac{\partial y^q}{\partial x^k}}_{\textbf{new term}}. \qquad (1.5) \end{aligned}$$

From the point of view of classical Riemannian geometry such definition of the high order invariant variation of y with terms $\Gamma(x)$, $\Gamma(y)$ and $\dfrac{\partial y_t^x}{\partial x}$ provides generalization of the classical covariant Riemannian derivative. Unlike all already existing torsion, polynomial connection and other generalizations of variation, defined primarily at point x, it depends not only on initial point of differentiation x, but also on behaviour of process at point y_t^x.

In terms of commutative diagrams approach, e. g. [10, 20], this definition does not proceed the traditional scheme for introduction of the high order tangent bundles over diffeomorphisms $M \ni x \to f(x) = y_t^x \in M$

$$\begin{array}{ccccccccc} M_x & \xrightarrow{T} & T_x M & \xrightarrow{T} & T_x^2 M & \xrightarrow{T} & \dots & \xrightarrow{T} & T_x^n M \\ \downarrow f & & \downarrow Tf & & \downarrow T^2 f & & & & \downarrow T^n f \\ M_{f(x)} & \xrightarrow{T} & T_{f(x)} M & \xrightarrow{T} & T_{f(x)}^2 M & \xrightarrow{T} & \dots & \xrightarrow{T} & T_{f(x)}^n M \end{array}$$

The additional term in (1.4), (1.5) makes an intermediate projection pr_H onto the corresponding horizontal component $H(T_{f(x)}^n M) \subset T_{f(x)}^n M$, isomorphic to the previous tangent space $H(T_{f(x)}^n M) \approx T_{f(x)}^{n-1} M$. Therefore new type variations (denoted by $\widetilde{T}^n f$ below)

$$\begin{array}{ccccccc}
T^n_x M & \xrightarrow{T} & T^{n+1}_x M & = & \cdots & = & T^{n+1}_x M \\
\downarrow \widetilde{T}^n f & & \downarrow T(\widetilde{T}^n f) & & & & \downarrow \widetilde{T}^{n+1} f \\
T_{f(x)} M & \xrightarrow{T} & T^2_{f(x)} M & \underset{pr_H}{\longrightarrow} & H(T^2_{f(x)} M) & \approx & T_{f(x)} M
\end{array}$$
$$\forall n \in I\!N$$

are recurrently forced to remain in $T_{f(x)}M$. To obtain the final picture, this commutative diagram should be transformed further to the tensorial bundles and covariant derivatives at point x. So one will be able to restrict attention to the changes of coordinates and corresponding covariances in (1.4), (1.5) and avoid the work with the differentiations in arbitrary directions on manifold, usually raised by functor T [10, 12].

Therefore for differential flows on manifolds we come to the more general concept, than the traditional tensor: we have to distinguish two different ways of dependence of tensor on point x: directly $u(x)$ and via process $u(y^x_t)$.

Definition 1.1. *Object $u^{(i/a)}_{(j/b)}$ is a generalized tensor if its coordinates*

$$u^{(i/a)}_{(j/b)} = u^{i_1 \dots i_p / a_1 \dots a_r}_{j_1 \dots j_q / b_1 \dots b_s}$$

form $T^{p,q}_x M$ tensor on multi-indexes $(i), (j)$ with respect to the local coordinates (x^k) and form $T^{r,s}M$ tensor on multi-indexes $(a), (b)$ with respect to the local coordinates (ϕ^m).

In other words, after the simultaneous change of local coordinate systems $(x^k) \to (x^{k'})$ and $(\phi^m) \to (\phi^{m'})$ one gets the transformation law

$$u^{(i/a)}_{(j/b)} = \frac{\partial x^{(i)}}{\partial x^{(i')}} \frac{\partial x^{(j')}}{\partial x^{(j)}} \frac{\partial \phi^{(a)}}{\partial \phi^{(a')}} \frac{\partial \phi^{(b')}}{\partial \phi^{(b)}} u^{(i'/a')}_{(j'/b')}, \tag{1.6}$$

where Jakobians have coordinate sense $\dfrac{\partial x^{(i)}}{\partial x^{(i')}} = \dfrac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \dfrac{\partial x^{i_p}}{\partial x^{i'_p}}$.

For differential flow, like $x \to y^x_t$, the new (ϕ^m) coordinates of the generalized tensor depend on coordinates (x^k). A simple example is provided by superpositions $u^{(a)}_{(b)}(y^x_t) v^{(\gamma)}_{(\delta)}(x)$, here the coordinate changes in domain $(x) \to (x')$ do not influence on $u^{(a)}_{(b)}(y^x_t)$, but one should write the additional Jacobians near $u^{(a)}_{(b)}(y^x_t)$ when making coordinate transformations $(y) \to (y')$. Another example is given by the first variation on initial data $\dfrac{\partial y^m(x,t)}{\partial x^k}$ – it represents a vector on index m in image (y_t^x) and covector field on index k in domain (x).

Next definition concretizes the idea of high order variations for the classical ordinary covariant ∇_k^x and stochastic derivatives D_z of generalized tensor. Recall that random function $F(\omega)$, defined on the Wiener space $\omega \in C_0(\mathbb{R}_+, \mathbb{R}^d)$, is *stochastically differentiable* [11, 18] in the direction of bounded continuous adapted to the canonical filtration process $z_t(\omega) \in \mathbb{R}^d$ if on the set of full measure there exists derivative

$$D_z F(\omega) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} F\bigg(\{\omega_t + \varepsilon \int_0^t z_s\, ds\}_{t\in\mathbb{R}^+}\bigg). \tag{1.7}$$

Definition 1.2. *The* **geometrically invariant ordinary** $\mathbb{V}_k^x$ *and* **stochastic** $\mathbb{D}_z$ *derivatives of generalized tensor are defined by*

$$\mathbb{V}_k^x u_{(j/b)}^{(i/a)} = \nabla_k^x u_{(j/b)}^{(i/a)} + \sum_{\rho\in(a)} \Gamma_{\gamma\ \delta}^{\ \rho}(\phi(x)) u_{(j/b)}^{(i/a)|_{\rho=\gamma}} \nabla_k^x \phi^\delta - \sum_{\rho\in(b)} \Gamma_{\rho\ \delta}^{\ \gamma}(\phi(x)) u_{(j/b)|_{\rho=\gamma}}^{(i/a)} \nabla_k^x \phi^\delta, \tag{1.8}$$

$$\mathbb{D}_z u_{(j/b)}^{(i/a)} = D_z u_{(j/b)}^{(i/a)} + \sum_{\rho\in(a)} \Gamma_{\gamma\ \delta}^{\ \rho}(\phi(x)) u_{(j/b)}^{(i/a)|_{\rho=\gamma}} D_z \phi^\delta - \sum_{\rho\in(b)} \Gamma_{\rho\ \delta}^{\ \gamma}(\phi(x)) u_{(j/b)|_{\rho=\gamma}}^{(i/a)} D_z \phi^\delta. \tag{1.9}$$

Above expression $(i/a)|_{\rho=\gamma}$ means that in the multi-index $(a) = (a_1, \ldots, a_\ell)$ instead of some index $\rho \in \{a_1, \ldots, a_\ell\}$ it is substituted index γ, which is a subject of summation on the same index γ of connection $\Gamma_{\gamma\ \delta}^{\ \rho}$ or $\Gamma_{\rho\ \delta}^{\ \gamma}$.

The additional terms in (1.8) and (1.9) make the resulting expression to be tensor with respect to the coordinates in image (ϕ^m). One may also note that the connection symbols above depend on parameters of image $(\phi(x))$, i.e. give nontraditional generalization of covariant differentiations. Moreover the dependence of y_t^x on the random Wiener space parameter $\omega \in C_0(\mathbb{R}_+, \mathbb{R}^d)$ does not evoke the problem of geometric invariance on parameter ω that takes values in $\mathbb{R}^d$. The tensorial character of new derivative can be easily checked [4, 6].

Theorem 1.1. *The invariant derivatives of generalized tensor are generalized tensor again, i. e. the tensor laws hold*

$$\mathbb{V}_k u^{(i/a)}_{(j/b)} = \frac{\partial x^{(i)}}{\partial x^{(i')}} \frac{\partial x^{(j')}}{\partial x^{(j)}} \frac{\partial \phi^{(a)}}{\partial \phi^{(a')}} \frac{\partial \phi^{(b')}}{\partial \phi^{(b)}} \frac{\partial x^{k'}}{\partial x^{k}} \mathbb{V}_{k'} u^{(i'/a')}_{(j'/b')},$$

$$\mathbb{D}_z u^{(i/a)}_{(j/b)} = \frac{\partial x^{(i)}}{\partial x^{(i')}} \frac{\partial x^{(j')}}{\partial x^{(j)}} \frac{\partial \phi^{(a)}}{\partial \phi^{(a')}} \frac{\partial \phi^{(b')}}{\partial \phi^{(b)}} \mathbb{D}_z u^{(i'/a')}_{(j'/b')}.$$

Proof. simply applies the definition of connection Γ and its transformation properties. □

2. Recurrent form of high order variational equations and place of curvature in the regularity problems.

Having in hands the correct procedure of differentiation of tensors of y^x_t, like defined by coefficients $A_0(y^x_t)$, $A_\sigma(y^x_t)$ of equation (1.2), we can ask about the place of curvature in the regularity properties.

The derivative on the initial data of the diffusion equation (1.2) can be obtained by direct differentiation and fulfills equation

$$\begin{aligned} \delta(\frac{\partial y^m}{\partial x^k}) &= \frac{\partial A^m_0(y)}{\partial x^k}\, dt + \frac{\partial A^m_\sigma(y)}{\partial x^k}\, \delta W^\sigma \\ &= (\mathbb{V}^x_k A^m_0(y) - \Gamma_{p\ q}^{\ m}(y) \frac{\partial y^p}{\partial x^k} A^q_0) dt \\ &+ (\mathbb{V}^x_k A^m_\sigma(y) - \Gamma_{p\ q}^{\ m}(y) \frac{\partial y^p}{\partial x^k} A^q_\sigma) \delta W^\sigma \\ &= -\Gamma_{p\ q}^{\ m}(y) \frac{\partial y^p}{\partial x^k} \delta y^q + \mathbb{V}^x_k A^m_0(y) dt + \mathbb{V}^x_k A^m_\sigma(y) \delta W^\sigma, \quad (2.1) \end{aligned}$$

where by adding and subtracting terms with connection we changed to the generalized derivatives of vector fields $A_\bullet(y^x_t)$ of image (y) and applied the symmetry of connection on lower indexes. We also used that near coefficients of connection stays the differential $\delta y = A_0(y)dt + A_\sigma(y)\delta W^\sigma$ of process (1.2).

Because stochastic derivative (1.7) has properties [11, 18]

$$D_z(f \circ (F_1, \dots, F_n)) = \sum_{j=1}^{n} (\partial_j f \circ (F_1, \dots, F_n)) D_z F_j,$$

$$D_z \int_0^t u(s)\, ds = \int_0^t D_z u(s)\, ds,$$

$$D_z \int_0^t u_\sigma(s)\delta W_s^\sigma = \int_0^t (D_z u_\sigma)\delta W_s^\sigma + \int_0^t u_\sigma(s) z_s^\sigma\, ds$$

we can also write the equation on the first order stochastic derivative

$$\begin{aligned}\delta(D_z y_t^m(\omega)) &= (D_z A_0^m(y) + A_\sigma^m(y) z^\sigma)dt + D_z A_\sigma^m(y)\delta W^\sigma \\ &= (\mathbb{D}_z A_0^m(y) - \Gamma_{p\,q}^{\,m}(y) D_z y^p A_0^q + A_\sigma^m(y) z^\sigma)dt \\ &\quad + (\mathbb{D}_z A_\sigma^m(y) - \Gamma_{p\,q}^{\,m}(y) D_z y^p A_\sigma^q)\delta W^\sigma \\ &= (\mathbb{D}_z A_0^m(y) + A_\sigma^m(y) z^\sigma)dt + \mathbb{D}_z A_\sigma^m(y)\delta W^\sigma - \Gamma_{p\,q}^{\,m} \mathbb{D}_z y^p \delta y^q, \quad (2.2)\end{aligned}$$

where we again add and subtract terms with connection to form the invariant stochastic derivatives (Definition 1.2) and separate the differential δy near connection.

Both equations (2.1) and (2.2) demonstrate that up to the parallel transition term with $\Gamma(y)$ the increments of first order variations are determined by invariant derivatives of coefficients. We take this observation as the recurrence base in the search for the high order variational equations.

Let X_γ^m denote some mixture of ordinary and stochastic variations, i.e.

$$X_\gamma^m = \mathcal{D}_\gamma y^m, \quad \mathcal{D}_\gamma = \mathcal{D}_{k_n} \dots \mathcal{D}_{k_1} \text{ for } \gamma = \{k_1, \dots, k_n\}, \tag{2.3}$$

where we introduced notation

$$\mathcal{D}_k = \begin{cases} \mathbb{V}_k, & \text{if } k \text{ is the "ordinary" index;} \\ \mathbb{D}_{z_k}, & \text{if } k \text{ is the "stochastic" index.} \end{cases} \tag{2.4}$$

Suppose that each equation on invariant variation has form, similar to the first order variational equations (2.1) or (2.2)

$$\delta(X_\gamma^m) = -\Gamma_{p\,q}^{\,m}(y) X_\gamma^p \delta y^q + M_{\gamma\,\sigma}^{\,m} \delta W^\sigma + N_\gamma^m dt \tag{2.5}$$

with some coefficients $M_{\gamma\,\sigma}^{\,m}$, N_γ^m. These coefficients we would like to determine by induction. For convenience of notations let us introduce an additional process $X_\emptyset$, that formally corresponds to the index $\gamma = \emptyset$ in (2.5)

$$\delta X_\emptyset^m = -\Gamma_{p\,q}^{\,m}(y) X_\emptyset^p \delta y^q + A_\sigma^m(y)\delta W^\sigma + A_0^m(y) dt.$$

The main result about the relation between different order variations gives next theorem, that reveals the role of curvature in regularity problems, see also Remark 2.1.

Theorem 2.1. *The relations between coefficients* $M_{\gamma\,\sigma}^{\,m}$, N_γ^m *for the process* X_γ^m *could be written in the terms of new type derivatives in the following compact form*

1) *recurrent base:*

$$M_{\emptyset\ \sigma}^{m} = A_\sigma^m(y_t^x), \quad N_\emptyset^m = A_0^m(y_t^x), \tag{2.6}$$

2) *recurrent step for $\gamma = \emptyset$ and for $\gamma \neq \emptyset$*

$$M_{\gamma\cup\{k\}\ \sigma}^{\ m} = \begin{cases} \mathcal{D}_k M_{\emptyset\ \sigma}^{\ m}, & \text{for } \gamma = \emptyset, \\ \mathcal{D}_k M_{\gamma\ \sigma}^{\ m} + R_{p\ \ell q}^{\ m} X_\gamma^p (\mathcal{D}_k y^\ell) A_\sigma^q, & \text{for } \gamma \neq \emptyset, \end{cases} \tag{2.7}$$

$$N_{\gamma\cup\{k\}}^m = \begin{cases} \mathcal{D}_k N_\emptyset^m + \lambda A_\sigma^m z_k^\sigma, & \text{for } \gamma = \emptyset, \\ \mathcal{D}_k N_\gamma^m + \lambda M_{\gamma\ \sigma}^{\ m} z_k^\sigma + R_{p\ \ell q}^{\ m} X_\gamma^p (\mathcal{D}_k y^\ell) A_\sigma^q, & \text{for } \gamma \neq \emptyset. \end{cases} \tag{2.8}$$

Here in (2.8) *the constant* $\lambda = 0$ *for the ordinary variation* $\mathcal{D}_k$*, i. e. corresponding process has form* $X_{\gamma\cup\{k\}} = \mathbb{V}_k X_\gamma$*, and* $\lambda = 1$ *for the stochastic variation* $\mathcal{D}_k$*, i. e. when process* $X_{\gamma\cup\{k\}} = \mathbb{D}_{z_k} X_\gamma$ *is constructed by stochastic differentiation of* X_γ*.*

As an obvious consequence of recurrent relations (2.6)–(2.8) one gets

Corollary 2.1. *The structure of coefficients* $M_{\gamma\ \sigma}^{m}$ *and* N_γ^m *is the following: for* $\gamma = \{k\}$, $|\gamma| = 1$

$$M_{k\ \sigma}^{\ m} = \mathcal{D}_k A_\sigma^m(y) = \nabla_\ell^y A_\sigma^m(y) \cdot \mathcal{D}_k y^\ell,$$
$$N_k^{\ m} = \mathcal{D}_k A_0^m(y) + \lambda A_\sigma^m z_k^\sigma,$$

where $\lambda = 0$ *for ordinary index* k*,* $\lambda = 1$ *for stochastic* k*.*

For higher order terms we have by (2.7), (2.8) *analogous representation*

$$M_{\gamma\ \sigma}^{\ m} = \nabla_\ell^y A_\sigma^m[\mathcal{D}_\gamma y^\ell] + \sum_{\beta_1\cup\ldots\cup\beta_s=\gamma,\ s\geq 2} L^1_{\beta_1,\ldots,\beta_s} \cdot \mathcal{D}_{\beta_1} y \ldots \mathcal{D}_{\beta_s} y, \tag{2.9}$$

$$\begin{aligned} N_\gamma^{\ m} = \nabla_\ell^y A_0^m[\mathcal{D}_\gamma y^\ell] &+ \sum_{\beta_1\cup\ldots\cup\beta_s=\gamma,\ s\geq 2} L^2_{\beta_1,\ldots,\beta_s} \cdot \mathcal{D}_{\beta_1} y \ldots \mathcal{D}_{\beta_s} y \\ &+ \lambda \sum_{\beta_1\cup\ldots\cup\beta_a\cup\varepsilon_1\cup\ldots\cup\varepsilon_b=\gamma} K_{\beta_1,\ldots,\beta_a,\varepsilon_1,\ldots,\varepsilon_b} \\ &\qquad \times \mathcal{D}_{\beta_1} y \ldots \mathcal{D}_{\beta_a} y \cdot \mathcal{D}_{\varepsilon_1\setminus k_1} z_{k_1} \ldots \mathcal{D}_{\varepsilon_b\setminus k_b} z_{k_b}. \end{aligned} \tag{2.10}$$

The summations in (2.9) runs on all nonintersecting subsets $\beta_1 \cup \ldots \cup \beta_a \cup \varepsilon_1 \cup \ldots \cup \varepsilon_b = \gamma$. Functions L^1, L^2 and K depend on A_0, A_σ, R and their covariant derivatives of order less than $|\gamma|$. In (2.10) notation $\mathcal{D}_{\varepsilon\setminus k}$ appears only for the case of stochastic derivatives $\mathcal{D} = \mathbb{D}_z$ on indexes

$k_1, \dots, k_b$ (then $\lambda = 1$) and means that in the set $\varepsilon = \{k_1, \dots, k_{|\varepsilon|}\}$ some point $k \in \varepsilon$ is removed. If the set ε consists of one point k, then the derivative $\mathcal{D}_{\varepsilon \setminus k}$ disappear and the corresponding summand is multiplied by z_k. If no stochastic indexes appear in set γ, then $\lambda = 0$.

Remark 2.1. Known approaches to define the variation to be covariant Riemannian, directional or stochastic derivative did not account the invariance on process y_t^x [10, 12, 18] and inevitably led to the growing number of non-invariant terms in the corresponding equations. Therefore it was principally hard to trace the influence of curvature in regular properties, especially in the noncompact manifold case.

The additional term with $\Gamma(y)$ in the Definition 1.2 of the new invariant derivative compactificates these non-invariant terms to the observable expressions with curvature. So it becomes possible to find the influence of curvature and non-linearities of diffusion equation on the any order regularity properties.

Proof of Theorem 2.1. The first step with $\gamma = \emptyset$ is already done in (2.1) and (2.2). We only demonstrate the recurrence step. The proof of stochastic $\mathbb{D}_z$ and ordinary $\mathbb{V}^x$ variations case is made separately.

1. Ordinary variation case. Below we omit, where possible, the dependence of connection Γ on variable y, however the dependence on x is always displayed precisely.

Let us simply substitute the definition of invariant derivative under Stratonovich integral

$$\int_0^t \delta(\mathbb{V}_k X_\gamma^m) = \int_0^t \delta\Big\{ \partial_k^x X_\gamma^m + \Gamma_{p\,q}^{\,m}(y) \frac{\partial y^p}{\partial x^k} X_\gamma^q - \sum_{s\in\gamma} \Gamma_{k\,s}^{\,h}(x) X_{\gamma|_{s=h}}^m \Big\}. \tag{2.11}$$

Now we use the properties of Stratonovich integrals

$$\int_0^t X\,\delta\Big(\int_0^t Y\,\delta Z\Big) = \int_0^t XY\,\delta Z,$$
$$\partial^x \int_0^t M\delta N = \int_0^t (\partial^x M)\delta N + \int_0^t M\delta(\partial^x N). \tag{2.12}$$

For the first term in (2.11) by inductive assumption (2.5) we find that

$$(2.11)_1 = \int_0^t \delta\Big(\partial_k^x \int_0^t \{-\Gamma_{p\,q}^{\,m} X_\gamma^p \,\delta y^q + M_{\gamma\,\sigma}^{\,m} \delta W^\sigma + N_\gamma^m dt\}\Big)$$

$$= -\int_0^t \frac{\partial \Gamma_{p\ q}^{\ m}}{\partial y^\ell}\frac{\partial y^\ell}{\partial x^k}\, X_\gamma^p\, \delta y^q - \int_0^t \Gamma_{p\ q}^{\ m}\, X_\gamma^p\, \delta\Big(\frac{\partial y^q}{\partial x^k}\Big) \tag{2.13}$$

$$-\int_0^t \Gamma_{p\ q}^{\ m}(\partial_k^x X_\gamma^p)\delta y^q + \int_0^t \{\partial_k^x M_{\gamma\ \sigma}^{\ m}\delta W^\sigma + \partial_k^x N_\gamma^m dt\}. \tag{2.14}$$

Let us rewrite the second term in (2.11) using Stratonovich-Itô formula

$$\delta\big(XYZ\big) = YZ\delta X + XZ\delta Y + XY\delta X$$

inductive assumption (2.5) and properties of Stratonovich integrals (2.12)

$$(2.11)_2 = \int_0^t \Gamma_{p\ q}^{\ m}\frac{\partial y^p}{\partial x^k}\delta(X_\gamma^q) + \int_0^t \Gamma_{p\ q}^{\ m}\, X_\gamma^q\, \delta\Big(\frac{\partial y^p}{\partial x^k}\Big) + \int_0^t \frac{\partial y^p}{\partial x^k}\, X_\gamma^q\, \delta \Gamma_{p\ q}^{\ m}(y)$$

$$= \int_0^t \Gamma_{p\ q}^{\ m}\frac{\partial y^p}{\partial x^k}\{-\Gamma_{\ell\ s}^{\ q}\, X_\gamma^\ell\, \delta y^s + M_{\gamma\ \sigma}^{\ q}\delta W^\sigma + N_\gamma^q dt\} \tag{2.15}$$

$$+\int_0^t \Gamma_{p\ q}^{\ m}\, X_\gamma^q\, \delta\Big(\frac{\partial y^p}{\partial x^k}\Big) + \int_0^t \frac{\partial y^p}{\partial x^k}\, X_\gamma^q\, \frac{\partial \Gamma_{p\ q}^{\ m}}{\partial y^\ell}\delta y^\ell. \tag{2.16}$$

With the last term in (2.11) we again apply the inductive assumption (2.5)

$$(2.11)_3 = -\sum_{s\in\gamma}\int_0^t \Gamma_{k\ s}^{\ h}(x)\{-\Gamma_{p\ q}^{\ m}(y) X^p_{\gamma|_{s=h}}\, \delta y^q$$

$$+ M_{\gamma|_{s=h}\ \sigma}^{\ m}\delta W^\sigma + N^m_{\gamma|_{s=h}} dt\}. \tag{2.17}$$

As the last preparation, we rewrite the first expression in (2.14) in terms of invariant derivative

$$(2.14)_1 = -\int_0^t \Gamma_{p\ q}^{\ m}\nabla_k X_\gamma^p \delta y^q + \int_0^t \Gamma_{p\ q}^{\ m}\Gamma_{\ell\ n}^{\ p}\frac{\partial y^\ell}{\partial x^k} X_\gamma^n\, \delta y^q$$

$$-\sum_{s\in\gamma}\int_0^t \Gamma_{p\ q}^{\ m}(y)\Gamma_{k\ s}^{\ h}(x) X^p_{\gamma|_{s=h}}\, \delta y^q. \tag{2.18}$$

Contracting the second expression in (2.13) with the first expression in (2.16), the second expression in (2.18) with the first expression in

(2.17) and forming from the second and third terms in (2.14), (2.15) and (2.17) the invariant derivatives of coefficients M and N, we have for the remaining terms

$$(2.11) = -\int_0^t \Gamma_{p\ q}^{\ m} \, \mathbb{V}_k X_\gamma^p \, \delta y^q + \int_0^t \{\mathbb{V}_k M_{\gamma\ \sigma}^{\ m} \delta W^\sigma + \mathbb{V}_k N_\gamma^m dt\} + \int_0^t \frac{\partial y^\ell}{\partial x^k} X_\gamma^p \, \delta y^q \Big\{ \frac{\partial \Gamma_{\ell\ p}^{\ m}(y)}{\partial y^q} - \frac{\partial \Gamma_{p\ q}^{\ m}(y)}{\partial y^\ell} + \Gamma_{s\ q}^{\ m}(y) \Gamma_{\ell\ p}^{\ s}(y) - \Gamma_{\ell\ s}^{\ m}(y) \Gamma_{p\ q}^{\ s}(y) \Big\}. \quad (2.19)$$

The terms in brackets in (2.19) appear correspondingly from the second term in (2.16), first term in (2.13), first term in (2.18) and first term in (2.15).

The expression in brackets $\{\ldots\}$ is a curvature tensor at point y_t^x, and we conclude

$$\int_0^t \delta(\mathbb{V}_k X_\gamma^m) = \int_0^t \Big\{ -\Gamma_{p\ q}^{\ m} (\mathbb{V}_k X_\gamma^p) \delta y^q + R_{p\ \ell q}^{\ m} X_\gamma^p \frac{\partial y^\ell}{\partial x^k} \delta y^q + \mathbb{V}_k M_{\gamma\ \sigma}^{\ m} \delta W^\sigma + \mathbb{V}_k N_\gamma^m dt \Big\},$$

i. e. the recurrence step for ordinary variation on initial data $\mathbb{V}^x$ is proved.

2. Stochastic variation. We proceed in a similar way.

By substitutions of inductive assumption and definition of invariant stochastic variation we get

$$\delta(I\!\!D_z X_\gamma^m) = \delta[D_z X_\gamma^m + \Gamma_{p\ q}^{\ m} X_\gamma^p D_z y^q]$$

$$= \delta\Bigg(D_z \Big[-\int_0^t \Gamma_{p\ q}^{\ m} X_\gamma^p \delta y^q + M_{\gamma\ \sigma}^{\ m} \delta W^\sigma + N_\gamma^m dt \Big] \Bigg) + \delta(\Gamma_{p\ q}^{\ m} X_\gamma^p D_z y^q)$$

$$= -\Gamma_{p\ q}^{\ m} D_z X_\gamma^p \delta y^q - \Gamma_{p\ q}^{\ m} X_\gamma^p \delta(D_z y^q) - \frac{\partial \Gamma_{p\ q}^{\ m}(y)}{\partial y^\ell} D_z y^\ell X_\gamma^p \delta y^q \quad (2.20)$$

$$+ (D_z M_{\gamma\ \sigma}^{\ m}) \delta W^\sigma + (M_{\gamma\ \sigma}^{\ m} z^\sigma + D_z N_\gamma^m) \, dt$$

$$+\frac{\partial\Gamma_{p\ q}^{\ m}(y)}{\partial y^\ell}X_\gamma^p D_z y^q \delta y^\ell + \Gamma_{p\ q}^{\ m} X_\gamma^p \delta(D_z y^q) + \Gamma_{p\ q}^{\ m} D_z y^q \delta(X_\gamma^p). \qquad (2.21)$$

Above we applied the properties of Stratonovich integrals (2.12), in particular that for stochastic derivative

$$D_z \int_0^t M\delta N = \int_0^t (D_z M)\delta N + \int_0^t M\delta(D_z N).$$

Next we contract the second terms in (2.20) and (2.21), extend derivatives D_z to the invariant derivative $I\!D_z$ and substitute inductive assumption (2.5) into the third term in (2.21) to find

$$\begin{aligned}\delta(I\!D_z X_\gamma^m) = &-\Gamma_{p\ q}^{\ m}(I\!D_z X_\gamma^p - \Gamma_{i\ j}^{\ p} X_\gamma^i D_z y^j)\delta y^q \\ &+ \left(\frac{\partial\Gamma_{p\ \ell}^{\ m}}{\partial y^q} - \frac{\partial\Gamma_{p\ q}^{\ m}}{\partial y^\ell}\right) X_\gamma^p D_z y^\ell \delta y^q \\ &+ (I\!D_z M_{\gamma\ \sigma}^{\ m} - \Gamma_{i\ j}^{\ m} M_{\gamma\ \sigma}^{\ i} D_z y^j)\delta W^\sigma \\ &+ (I\!D_z N_\gamma^m - \Gamma_{i\ j}^{\ m} N_\gamma^i D_z y^j)dt + M_{\gamma\ \sigma}^{\ m} z^\sigma dt \\ &+ \Gamma_{p\ q}^{\ m} D_z y^q[-\Gamma_{i\ j}^{\ p} D_z y^i \delta y^j + M_{\gamma\ \sigma}^{\ p}\delta W^\sigma + N_\gamma^p dt] \\ = &-\Gamma_{p\ q}^{\ m} I\!D_z X_\gamma^p \delta y^q + I\!D_z M_{\gamma\ \sigma}^{\ m}\delta W^\sigma \\ &+ (I\!D_z N_\gamma^m + M_{\gamma\ \sigma}^{\ m} z^\sigma)dt + R_{p\ \ell q}^{\ m} X_\gamma^p D_z y^\ell \delta y^q.\end{aligned}$$

Like in the step 1 of proof, one comes to the curvature tensor. □

3. Smooth and smoothing representations of semigroup derivatives.

Now we ask about the role that plays new type ordinary and stochastic variations in the regularity properties. In [4] it was proved representation

$$(\nabla^x)^i P_t f(x) = \sum_{\substack{i_1+\ldots+i_s=i,\\ s=1,\ldots,i}} \mathbf{E}\langle(\nabla^y)^s f(y_t^x), (\mathbb{W}^x)^{i_1} y_t^x \otimes\ldots\otimes(\mathbb{W}^x)^{i_s} y_t^x\rangle. \quad (3.1)$$

Notation (3.1) means:

$$\nabla_{k_i}^x \ldots \nabla_{k_1}^x P_t f(x) = \sum_{\gamma_1\cup\ldots\cup\gamma_s=\{k_1,\ldots,k_i\}} \mathbf{E}\ \nabla_{j_s}^y \ldots \nabla_{j_1}^y f(y_t^x) \mathbb{W}_{\gamma_1}^x y^{j_1} \ldots \mathbb{W}_{\gamma_s}^x y^{j_s}. \quad (3.2)$$

We also denoted $\mathbb{V}^x y_t^x$ for the first order variation $\dfrac{\partial y_t^x}{\partial x}$. Further we will use both coordinate (3.1) and coordinate-free (3.2) notations.

To obtain raise of smoothness representations, let us note that due to the chain rules the principal parts of equations on the first order ordinary and stochastic variations (2.1) and (2.2) coincide. Therefore, by variation of constants method we can find a special stochastic direction $\widetilde{z}_k$ such that

$$D_{\widetilde{z}_k} y_t^m = t \mathbb{V}_k y_t^m \tag{3.3}$$

A simple calculation of the Stratonovich differential by (2.1)

$$\delta\Big(t \frac{\partial y_t^m}{\partial x^k}\Big) = \frac{\partial y_t^m}{\partial x^k} dt - t \Big\{ \Gamma_{p\ q}^{\ m} \frac{\partial y^p}{\partial x^k} \delta y^q + \mathbb{V}_k A_\sigma^m(y) \delta W^\sigma + \mathbb{V}_k A_0^m(y) dt \Big\}$$

with further substitution of (3.3) and therefore of $\mathbb{D}_{\widetilde{z}_k} A_\sigma^m(y_t^x) = t \mathbb{V}_k A_\sigma^m(y_t^x)$ permits to transform the above relation to form (2.2) if we choose the inhomogeneous part so that

$$\frac{\partial y_t^m}{\partial x^k} = A_\sigma^m(y_t^x) \widetilde{z}_k^\sigma(t). \tag{3.4}$$

Henceforth we use notation $\widetilde{z}$ for a special choice of stochastic direction (3.4).

Similar to [7], choice (3.3) leads to the first order raise of smoothness representation:

$$\begin{aligned} \nabla_k P_t f(x) = \mathbf{E} \nabla_m^y f(y_t^x) \cdot \frac{\partial y_t^m}{\partial x^k} &= \frac{1}{t} \mathbf{E} \nabla_m^y f(y_t^x) \cdot D_{\widetilde{z}_k} y_t^m \\ &= \frac{1}{t} \mathbf{E} D_{\widetilde{z}_k} f(y_t^x) = \frac{1}{t} \mathbf{E} f(y_t^x) \int_0^t \widetilde{z}_k^\sigma dW^\sigma, \end{aligned} \tag{3.5}$$

here we used the integration by parts characterization of Wiener measure [11, 18]:

$$\mathbf{E} D_z F = \mathbf{E} F \int_0^\infty z_s^\sigma dW^\sigma.$$

Thus for non-exploding stochastic integral $\int_0^t \widetilde{z}_k^\sigma dW^\sigma$ and continuous function f in the r.h.s. above the semigroup $P_t f$ becomes one time differentiable on x for all $t > 0$.

To find higher order representations we need the invariant form of integration-by-parts formula for generalized tensors (Def. 1.1).

Theorem 3.1. *For tensors $F^{(a)}_{(b)}$ and $G^{(b)}_{(a)}$, depending on the process y^x_t the integration by parts formula holds*

$$\mathbf{E}(I\!D_z F^{(a)}_{(b)})\, G^{(b)}_{(a)} = -\mathbf{E}\, F^{(a)}_{(b)}\, I\!D_z G^{(b)}_{(a)} + \mathbf{E}\, F^{(a)}_{(b)}\, G^{(b)}_{(a)} \int\limits_0^\infty z^\sigma dW^\sigma. \tag{3.6}$$

Here the summation on repeating multi-indexes is implemented.

Proof. Let us take two tensors, depending on (y^x_t). By Definition 1.9

$$\begin{aligned}
\mathbf{E}(I\!D_z F^{(a)}_{(b)}) G^{(b)}_{(a)} &= \mathbf{E}\Big\{ D_z F^{(a)}_{(b)} + \sum_{s\in(a)} \Gamma_p{}^s{}_q (D_z y^q) F^{(a)|_{s=p}}_{(\beta)} \\
&\qquad - \sum_{s\in(b)} \Gamma_s{}^p{}_q (D_z y^q) F^{(a)}_{(b)|_{s=p}} \Big\} G^{(b)}_{(a)} \\
&= \mathbf{E}\Big\{ -F^{(a)}_{(b)} D_z G^{(b)}_{(a)} + F^{(a)}_{(b)} G^{(b)}_{(a)} \int\limits_0^\infty z^\sigma dW^\sigma \\
&+ \sum_{s\in(a)} \Gamma_s{}^p{}_q (D_z y^q) F^{(a)}_{(b)} G^{(b)}_{(a)|_{s=p}} - \sum_{s\in(b)} \Gamma_p{}^s{}_q (D_z y^q) F^{(a)}_{(b)} G^{(b)|_{s=p}}_{(a)} \Big\}.
\end{aligned}$$

This implies formula (3.6). Above we used integration by parts for Wiener measure and redenoted indexes p and s. ☐

Now we can find high order raise of smoothness representations, that give $P_t : C^n \to C^{n+m}$ for $n = 0,\ m \in I\!N$. This result generalizes [7] to the manifold setting.

Let us introduce the following notation:

$$Y_k = t\mathbb{V}^x_k - \widetilde{I\!D}_k + \int\limits_0^t \widetilde{z}^\sigma_k dW^\sigma,$$

where

$$\widetilde{I\!D}_k = I\!D_{\widetilde{z}_k} \tag{3.7}$$

for the invariant stochastic derivative in particular direction $\widetilde{z}_k$ chosen in (3.4).

Theorem 3.2. *High order covariant derivative of semigroup P_t permits representation:*

$$\nabla^x_\gamma P_t f(x) = \frac{1}{t^{|\gamma|}} \mathbf{E} f(y^x_t) Y_\gamma\, 1. \tag{3.8}$$

Above $Y_\gamma = Y_{k_n} \dots Y_{k_1}$ for a set $\gamma = \{k_1, \dots, k_n\}$.

Proof. proceeds in analogue to [7] with use of invariant derivatives $\mathbb{\nabla}^x_k$, $\widetilde{I\!\!D}_k$ and integration by parts (3.6). □

4. Nonlinear estimate on variations and manifold form of dissipativity and coercitivity conditions.

From representations (3.1) and (3.8) we see that the smooth and smoothing properties of semigroups require to study the new type variations of processes X_γ (2.3), when stochastic derivatives are taken in particular directions $\widetilde{z}_k$ (3.3),(3.4).

To proceed further we use an important observation about symmetries of high order derivatives of nonlinear functions, e. g. [2]–[3]: in the high order differentials

$$d^n F(y) = F'(y) d^n y + \sum_{\substack{j_1+\dots+j_i=n, \\ i=2,\ n-1}} F^{(i)}(y) d^{j_1} y \dots d^{j_i} y + F^{(n)}(y)(dy)^n$$

simultaneously arise n^{th} order differential and 1^{st} order differential in n^{th} power, thus the differentials are comparable: $d^n y \sim (dy)^n$. This symmetry is also manifested in the intermediate terms $d^{j_1} y \dots d^{j_i} y \sim (dy)^{j_1+\dots+j_i} \sim (dy)^n$ and hold for any n.

An additional property of the invariant variations

$$\mathcal{D}\big(u^{(a)}_{(b)} \circ y^x_t\big) = \big(\nabla^y_m u^{(a)}_{(b)}\big) \circ (y^x_t) \cdot \mathcal{D} y^m, \tag{4.1}$$

that could be easily checked from definition, permits us to look closely on the variational equations (2.5). We see from the chain rule (4.1) and Corollary 2.1 that the high order variation δX^m_γ in the l. h. s. is *proportional* to the product of lower order variations $X_{\beta_1} \dots X_{\beta_i}$, $\beta_1 \cup \dots \cup \beta_i = \gamma$, arising in the r. h. s. of (2.5). More precisely, accounting property (3.3) of $\widetilde{z}_k$ we come to the chain of symmetries

$$(\mathbb{\nabla}^x)^i y^x_t \approx (\mathbb{\nabla}^x y^x_t)^i \approx \frac{1}{t^i} (\widetilde{I\!\!D} y^x_t)^i \approx \frac{1}{t^i} \widetilde{I\!\!D}^i y^x_t \approx \frac{1}{t^{|\beta|}} \mathcal{D}^{\alpha \cup \beta} y^x_t, \tag{4.2}$$

where $\alpha \cup \beta = \{1, \dots, i\}$ represents subdivision of set $\{1, \dots, i\}$ on the part, that corresponds to the ordinary differentiation of order α and stochastic differentiation of order β. Henceforth we introduce notation

$$\mathcal{D}_k = \begin{cases} \mathbb{\nabla}_k, & \text{if } k \text{ is the "ordinary" index,} \\ \widetilde{I\!\!D}_k, & \text{if } k \text{ is the "stochastic" index,} \end{cases}$$

where $\widetilde{I\!D}_k$ means the stochastic derivative in special direction $\widetilde{z}_k$ (3.4). Corresponding expression $I\!D^{\alpha\cup\beta} y_t^x$ represents the coordinate free notation for the high order variation, where multi-index α corresponds to ordinary differentiation, and β to stochastic one, $\alpha \cup \beta = \{1, \dots, i\}$ for some i. The upper indexes in $I\!D^\gamma$ mean the coordinate-free notations.

Consider the nonlinear expression that reflects symmetry (4.2)

$$Q_n^{n'}(y,t) = \sum_{\substack{\alpha\cup\beta=\{1,\dots,i\},\\ |\alpha|\le n,\ |\beta|\le n'}} \mathbf{E} p_i(\rho^2(y_t^x, o)) \, \| \frac{1}{t^{|\beta|}} I\!D^{\alpha\cup\beta} y_t^x \|^{r/i}, \ \ r \ge 2(n+n') \tag{4.3}$$

and simultaneously represents some quasi-norm on the high order regularity.

Remark that factor r/i makes expression (4.3) to be homogeneous with respect to the nonlinear symmetry (4.2). Above o is some fixed point of manifold, that plays a role of reference point, analogous to 0 in $I\!R^d$.

Because the symmetry (4.2) manifests also in the variational kernels in representations of semigroups derivatives (3.1) and (3.8), the knowledge of these symmetries and corresponding estimates on nonlinear expressions (4.3) becomes essential for the study of regularity properties [2, 3].

Theorem 4.1. *Suppose that the following conditions on the coefficients of equation* (1.2) *and geometry of manifold hold*

- **dissipativity***: $\exists\, o \in M$ such that $\forall\, C \in I\!R^1$ $\exists\, K_C \in I\!R^1$ such that $\forall\, x \in M$*

$$\langle \widetilde{A_0}(x), \nabla^x \rho_M^2(x,o)\rangle + C \sum_{\sigma=1}^d \|A_\sigma(x)\|^2 \le K_C(1 + \rho_M^2(x,o)), \tag{4.4}$$

- **differential coercitivity***: $\forall\, C,\ C' \in I\!R^1$ $\exists\, K_C \in I\!R^1$ such that $\forall\, x \in M$, $\forall\, h \in T_x M$*

$$\langle \nabla \widetilde{A_0}(x)[h], h\rangle + C \sum_{\sigma=1}^d \|\nabla A_\sigma(x)[h]\|^2 + C' \sum_{\sigma=1}^d \langle R_x(A_\sigma(x), h) A_\sigma(x), h\rangle \le K_C \|h\|^2, \tag{4.5}$$

where $\widetilde{A_0} = A_0 + \frac{1}{2}\sum_{\sigma=1}^{d} \nabla_{A_\sigma} A_\sigma$ *and* $R(A,h)A = R_{p\ \ell q}^{\ m} A^p A^\ell h^q$. *Henceforth notation* $\nabla H[h]$ *means the* directional covariant derivative, *defined by*

$$(\nabla H(x)[h])^i = \nabla_j H^i(x) \cdot h^j, \tag{4.6}$$

- **nonlinear behaviour of coefficients and curvature**: *for any* n *there are constants* $\mathbf{k}_\bullet$ *such that for all* $j = 1, \ldots, n$ *and* $\forall\, x \in M$

$$\begin{aligned} \|(\nabla)^j \widetilde{A_0}(x)\| &\le (1 + \rho_M(x,o))^{\mathbf{k}_0}, \\ \|(\nabla)^j A_\sigma(x)\| &\le (1 + \rho_M(x,o))^{\mathbf{k}_\sigma}, \\ \|(\nabla)^j R(x)\| &\le (1 + \rho_M(x,o))^{\mathbf{k}_R} \end{aligned} \tag{4.7}$$

$$\exists\, \mathbf{k}_1 \text{ such that } \inf \frac{\|A^\sigma(x)\|}{(1+\rho^2(x,o))^{\mathbf{k}_1}} > 0. \tag{4.8}$$

Then is some $\mathbf{k} = \mathbf{k}(\mathbf{k}_0, \mathbf{k}_1, \mathbf{k}_\sigma, \mathbf{k}_R)$ *such that if in* (4.3) *monotone polynomials* $p_j \ge 1$ *are hierarchied by*

$$\forall\, j_1 + \ldots + j_s = i \le n \quad [p_i(|\cdot|^2)]^i (1 + |\cdot|^2)^{r\mathbf{k}} \le [p_{j_1}(|\cdot|^2)]^{j_1} \ldots [p_{j_s}(|\cdot|^2)]^{j_s} \tag{4.9}$$

the nonlinear estimate on ordinary and stochastic variations holds

$$\exists\, K_{\mathbf{k}} \quad \forall\, t \ge 0 \quad Q_n^{n'}(y,t) \le e^{K_{\mathbf{k}} t} Q^0_{n+n'}(y,0). \tag{4.10}$$

Remark 4.1. Conditions (4.4)–(4.5) generalize the classical Krylov-Rosovskii-Pardoux conditions [16, 19] from the linear base space to manifold. They relate the nonlinearities of diffusion equation with the geometric properties of manifold, without traditional separation of geometry.

Proof of Theorem 4.1. Let us first explain the idea of proof, that develops [3]. One should take the differential of nonlinear term in (4.3) and estimate it by the same expression from above to get the exponential estimate (4.10). For this we first find the recurrent on γ form for the differential of norm of solutions to (2.5). Collecting the representations of these differentials, we can then prove nonlinear estimate (4.10).

As an important moment of the proof one should prepare the estimates on differential of $\rho_M^2(y_t^x, o)$ in a way, that avoids the complicate formulas of geodesic deviations and introduction of Jacobi fields.

Theorem 4.2 ([9]). *Suppose that the generalized dissipativity and coercitivity conditions* (4.4)–(4.5) *hold.*

Then there is constant K such that

$$\Big\{A_0^1 + A_0^2 + \frac{1}{2}\sum_{\sigma=1}^{d}(A_\sigma^1 + A_\sigma^y)^2\Big\}\rho^2(x,y) \le K\rho^2(x,y). \tag{4.11}$$

Similarly $\forall\, C\ \exists\, K_C$ such that

$$\mathcal{L}^1\rho^2(x,y) + C\sum_{\sigma=1}^{d}\frac{(A_\sigma^1\rho^2(x,y))^2}{\rho^2(x,y)} \le K(1+\rho^2(x,y)). \tag{4.12}$$

Operator $\mathcal{L}$ (1.4) *is generator of diffusion* (1.2).

The proof of this result may be found in [9].

Step 1. Consider one of the summands in the expression (4.3), corresponding to set $\alpha\cup\beta = \{1,\dots,i\}$ without factor $\frac{1}{t^{|\beta|}}$. We put for convenience $\frac{r}{i} = 2q$ and use notation $\rho(y_s^t)$ instead of $\rho(y_t^s, o)$. By Itô formula

$$h(t) = \mathbf{E}p_i(\rho^2(y_t^x))\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2q} = h(0)$$

$$+\mathbf{E}\int_0^t p_i(\rho^2(y_t^x))\, d\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2q} \tag{4.13}$$

$$+\mathbf{E}\int_0^t \|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2q} dp_i(\rho^2(y_t^x)) \tag{4.14}$$

$$+\frac{1}{2}\mathbf{E}\int_0^t d[p_i(\rho^2(y_t^x)), \|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2q}]. \tag{4.15}$$

In the last term the brackets $[X,Y]$ mean the quadratic variation of processes X and Y. Proceeding further and applying Itô formula we have

$$(4.13) = q\int_0^t \mathbf{E}p_i(\rho^2(y_t^x))\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2(q-1)}d\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2$$

$$+\frac{1}{2}q(q-1)\int_0^t \mathbf{E}p_i(\rho^2(y_t^x))\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2(q-2)}d\big[\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2, \|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2\big]; \tag{4.16}$$

$$(4.14) = \int_0^t \mathbf{E}\|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2q} p_i'(\rho^2(y_t^x))d\rho^2(y_t^x) + \frac{1}{2}\int_0^t \mathbf{E}\|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2q} p_i''(\rho^2(y_t^x))d[\,\rho^2(y_t^x),\rho^2(y_t^x)\,]; \quad (4.17)$$

$$(4.15) = \frac{1}{2}\int_0^t \mathbf{E} p_i'(\rho^2(y_t^x))\|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2(q-1)} d\big[\rho^2(y_t^x), \|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^2\big]. \quad (4.18)$$

Step 2. To estimate (4.13) we use the following Theorem, proved in [6].

Theorem 4.3. *The differential of norm of process X_γ^m (2.5) has form*

$$d\|X_\gamma^m\|^2 = g^{\gamma\varepsilon}(x)\,\{\,g_{mn}(X_\gamma^m M_{\varepsilon\ \sigma}^{\ n} + X_\varepsilon^n M_{\gamma\ \sigma}^{\ m})dW^\sigma + g_{mn}(X_\gamma^m N_\varepsilon^n + X_\varepsilon^n N_\gamma^m + M_{\gamma\ \sigma}^{\ m} M_{\varepsilon\ \sigma}^{\ n})dt + \frac{1}{2} g_{mn}(X_\gamma^m P_\varepsilon^n + X_\varepsilon^n P_\gamma^m)dt\}, \quad (4.19)$$

where expressions P_γ^m are defined in the following recurrent way

$$P_k^m = \mathcal{D}_k(\nabla_{A_\sigma} A_\sigma^m) + R_{p\ \ell q}^{\ m} A_\sigma^p A_\sigma^q (\mathcal{D}_k y^\ell), \quad (4.20)$$

$$P_{\gamma\cup\{k\}}^m = \mathcal{D}_k P_\gamma^m + 2R_{p\ \ell q}^{\ m} M_{\gamma\ \sigma}^{\ p}(\mathcal{D}_k y^\ell)A_\sigma^q + (\nabla_s R_{p\ \ell q}^{\ m})X_\gamma^p(\mathcal{D}_k y^\ell)A_\sigma^q A_\sigma^s + R_{p\ \ell q}^{\ m} X_\gamma^p(\mathcal{D}_k A_\sigma^\ell)A_\sigma^q + R_{p\ \ell q}^{\ m} X_\gamma^p(\mathcal{D}_k y^\ell)(\nabla_{A_\sigma} A_\sigma). \quad (4.21)$$

Now, Corollary 2.1 permits to find the structure of coefficients P_γ^m. For $|\gamma| = 1$ and $X_k^m = \mathcal{D}_k y^m$ from (4.20) we have

$$P_k^m = \nabla_\ell^y \nabla_{A_\sigma} A_\sigma^m \cdot \mathcal{D}_k y^\ell + R(A_\sigma, \mathcal{D}_k y)A_\sigma.$$

Due to in (4.21) $P_{\gamma\cup\{k\}}^m = \mathcal{D}_k P_\gamma^m + \ldots$ (coefficient with highest order differentiation of y: $\mathcal{D}_\gamma y$), the coefficients P_γ^m permit representation

$$P_\gamma = \nabla_\ell \nabla_{A_\sigma} A_\sigma \cdot \mathcal{D}_\gamma y^\ell + R(A_\sigma, \mathcal{D}_\gamma y)A_\sigma + \sum_{\delta_1\cup\ldots\cup\delta_s=\gamma,\ s\geq 2} L_{\delta_1,\ldots,\delta_s} \cdot \mathcal{D}_{\delta_1} y \ldots \mathcal{D}_{\delta_s} y,$$

with coefficients $L_{\delta_1,\dots,\delta_s}$ depending on A_0, A_σ, R and their covariant derivatives. This expression contains symmetries (4.2) in dependence on lower order variations $\mathcal{D}_\delta y$. Substituting this into (4.19) we have

$$d\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2 = 2\langle \mathcal{D}^{\alpha\cup\beta}y, \nabla^y_\ell A_\sigma[\mathcal{D}^{\alpha\cup\beta}y^\ell]\rangle_{T_x^{(0,i)}\otimes T_y} dW^\sigma \tag{4.22}$$

$$+2\langle \mathcal{D}^{\alpha\cup\beta}y, \nabla^y_\ell \widetilde{A_0}[\mathcal{D}^{\alpha\cup\beta}y^\ell]\rangle_{T_x^{(0,i)}\otimes T_y} dt \tag{4.23}$$

$$+\sum_{\sigma=1}^{d} \|\nabla A_\sigma[\mathcal{D}^{\alpha\cup\beta}y]\,\|^2_{T_x^{(0,i)}\otimes T_y} dt \tag{4.24}$$

$$+\sum_{\sigma=1}^{d} \langle R(A_\sigma, \mathcal{D}^{\alpha\cup\beta}y)A_\sigma, \mathcal{D}^{\alpha\cup\beta}y\rangle_{T_x^{(0,i)}\otimes T_y} dt \tag{4.25}$$

$$+\sum_{\delta_1+\ldots+\delta_s=\alpha\cup\beta,\ s\geq 2} L^3_{\delta_1,\dots,\delta_s,\sigma}\langle \mathcal{D}^{\alpha\cup\beta}y, \mathcal{D}^{\delta_1}y\ldots\mathcal{D}^{\delta_s}y\rangle\, dW^\sigma \tag{4.26}$$

$$+\sum_{\delta_1+\ldots+\delta_s=\alpha\cup\beta,\ s\geq 2} L^4_{\delta_1,\dots,\delta_s,\sigma}\langle \mathcal{D}^{\alpha\cup\beta}y, \mathcal{D}^{\delta_1}y\ldots\mathcal{D}^{\delta_s}y\rangle\, dt \tag{4.27}$$

$$+\lambda\sum_{\substack{\beta_1,\dots,\beta_a,\\ \varepsilon_1,\dots,\varepsilon_b}} K^1_{\beta_1,\dots,\beta_a,\varepsilon_1,\dots,\varepsilon_b}\langle \mathcal{D}^{\alpha\cup\beta}y, \mathcal{D}^{\beta_1}y\ldots\mathcal{D}^{\beta_a}y\cdot\mathcal{D}^{\varepsilon_1\setminus k_1}\widetilde{z}_{k_1}\ldots\mathcal{D}^{\varepsilon_b\setminus k_b}\widetilde{z}_{k_b}\rangle dt. \tag{4.28}$$

In formulas (4.26) and (4.27) sets δ contain both ordinary and stochastic indexes. Summation in (4.28) runs on all subdivision of set $\{1,\dots,i\}$ on nonintersecting subsets $\beta_1\cup\ldots\cup\beta_a\cup\varepsilon_1\cup\ldots\cup\varepsilon_b = \{1,\dots,i\}$ see (2.10). Like before the coefficients $L^3,\ L^4,\ K^1$ depend in a polynomial way (4.7) on covariant derivatives of $A_0,\ A_\sigma,\ R$ and display symmetry (4.2) on variations and $\mathcal{D}^{\varepsilon\setminus k}\widetilde{z}_k$ terms (on both index of differentiation and lower index at $\widetilde{z}$.).

From (4.22), (4.26) it follows that expression for quadratic variation of $\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2$ can be estimated as follows:

$$d\left[\,\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2, \|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2\,\right] \leq 4\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2\cdot\sum_{\sigma=1}^{d}\|\nabla^y A_\sigma[\mathcal{D}^{\alpha\cup\beta}y_t^x]\|^2 dt \tag{4.29}$$

$$+\sum_{\substack{\delta_1+\ldots+\delta_s=\alpha\cup\beta,\\ s\geq 2}} L^5_{\delta_1,\dots,\delta_s,\sigma}\,\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2\cdot\|\langle \mathcal{D}^{\alpha\cup\beta}y_t^x, \mathcal{D}^{\delta_1}y_t^x\ldots\mathcal{D}^{\delta_s}y_t^x\rangle\|\, dt \tag{4.30}$$

$$+\sum_{\delta_1+\ldots+\delta_s=\alpha\cup\beta, s\geq 2} L^6_{\delta_1,\dots,\delta_s,\sigma}\,\|\langle \mathcal{D}^{\alpha\cup\beta}y_t^x, \mathcal{D}^{\delta_1}y_t^x\ldots\mathcal{D}^{\delta_s}y_t^x\rangle\|^2 dt. \tag{4.31}$$

Collecting terms, which appear from (4.23), (4.24), (4.25) and (4.29) we have

$$(4.13) \le 4q \int_0^t \mathbf{E} p_i(\rho^2(y_t^x)) \|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2(q-1)} \langle \mathcal{D}^{\alpha\cup\beta} y, \nabla_\ell^y \widetilde{A_0}[\mathcal{D}^{\alpha\cup\beta} y^\ell]\rangle \, dt \tag{4.32}$$

$$+(2q+4) \int_0^t \mathbf{E} p_i(\rho^2(y_t^x)) \|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2(q-1)} \sum_{\sigma=1}^d \|\nabla A_\sigma[\mathcal{D}^{\alpha\cup\beta} y]\,\|^2 dt \tag{4.33}$$

$$+2q \int_0^t \mathbf{E} p_i(\rho^2(y_t^x)) \|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2(q-1)} \sum_{\sigma=1}^d \langle R(A_\sigma, \mathcal{D}^{\alpha\cup\beta} y) A_\sigma, \mathcal{D}^{\alpha\cup\beta} y\rangle \, dt \tag{4.34}$$

$$+I_1 + I_2 + I_3, \tag{4.35}$$

where

$$I_1 = \frac{1}{2} q(q+1) \int_0^t \sum_{\delta_1 \ldots \delta_s} \mathbf{E}\, L^7_{\delta_1,\ldots,\delta_s} p_i(\rho^2(y_t^x)) \times \|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2q-1} \|\mathcal{D}^{\delta_1} y_t^x \ldots \mathcal{D}^{\delta_s} y_t^x\| \, dt; \tag{4.36}$$

$$I_2 = \lambda q \int_0^t \sum_{\substack{\beta_1\ldots\beta_a,\\ \varepsilon_1\ldots\varepsilon_b}} \mathbf{E}\, K^1_{\beta_1,\ldots,\beta_a,\varepsilon_1,\ldots,\varepsilon_b}\, p_i(\rho^2(y_t^x)) \|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2q-1} \times \|\mathcal{D}^{\beta_1} y_t^x \ldots \mathcal{D}^{\beta_a} y_t^x \cdot \mathcal{D}^{\varepsilon_1 \setminus k_1} \widetilde{z}_{k_1} \ldots \mathcal{D}^{\varepsilon_b \setminus k_b} \widetilde{z}_{k_b}\| \, dt; \tag{4.37}$$

$$I_3 = \frac{1}{2} q(q-1) \int_0^t \sum_{\delta_1 \ldots \delta_s} \mathbf{E}\, L^6_{\delta_1,\ldots,\delta_s,\sigma}\, p_i(\rho^2(y_t^x)) \times \|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2q-2} \|\mathcal{D}^{\delta_1} y_t^x \ldots \mathcal{D}^{\delta_s} y_t^x\|^2 dt. \tag{4.38}$$

Step 3. Terms (4.17) and (4.18) are estimated in a similar way. Term (4.17) is transformed by monotonicity of polynomials $p_i(\cdot)$. Applying Itô formula to $\rho^2(y_t^x)$

$$\rho^2(y_t^x) = \rho^2(x) + \sum_{\sigma=1}^d \int_0^t [(A^1_\sigma + A^2_\sigma)\rho^2](y_s^x) \, dW_s^\sigma$$

$$+\int_0^t \{A_0^1 + A_0^2 + \frac{1}{2}\sum_{\sigma=1}^d (A_\sigma^1 + A_\sigma^2)^2\}\rho^2(y_s^x)\,ds \quad (4.39)$$

we can continue

$$(4.17) = \int_0^t \mathbf{E}\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2q}\Big\{p_i'(\rho^2(y_t^x))\mathcal{L}\rho^2(y_t^x) + \frac{1}{2}p_i''(\rho^2(y_t^x))\rho^2(y_t^x)\frac{1}{\rho^2(y_t^x)}\sum_{\sigma=1}^d (A_\sigma\rho^2(y_t^x))^2\Big\}\,dt$$

$$\leq \int_0^t \mathbf{E}\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2q}p_i'(\rho^2(y_t^x))\Big\{\mathcal{L}\rho^2(y_t^x) + \frac{C}{\rho^2(y_t^x)}\sum_{\sigma=1}^d \|A_\sigma\rho^2(y_t^x)\|^2\Big\}\,dt. \quad (4.40)$$

Due to representation (4.22)–(4.28) of $\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2$ and Itô formula for metric $\rho(y_t^x)$ (4.39) we obtain

$$d\big[\rho^2(y_t^x), \|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^2\big] = 2\sum_{\sigma=1}^d A_\sigma\rho^2(y_t^x)\big\langle\mathcal{D}^{\alpha\cup\beta}y_t^x, \{\nabla A_\sigma[\mathcal{D}^{\alpha\cup\beta}y_t^x] + \sum_{\delta_1,\dots,\delta_s} L^3_{\delta_1,\dots,\delta_s,\sigma}\mathcal{D}^{\delta_1}y_t^x\dots\mathcal{D}^{\delta_s}y_t^x\}\big\rangle dt.$$

Using inequality $2a\|x\|\,\|y\| \leq \frac{a^2\|x\|^2}{\rho^2} + \|y\|^2\rho^2$ we estimate

$$(4.18) \leq \int_0^t \mathbf{E}\,p_i'(\rho^2(y_t^x))\,\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2(q-1)} \times\Big|\sum_{\sigma=1}^d A_\sigma\rho^2(y_t^x)\cdot\big\langle\mathcal{D}^{\alpha\cup\beta}y_t^x, \nabla A_\sigma[\mathcal{D}^{\alpha\cup\beta}y_t^x]\big\rangle\Big|dt$$

$$+\int_0^t \mathbf{E}\,p_i'(\rho^2(y_t^x))\,\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2(q-1)} \times\Big|\sum_{\sigma=1}^d A_\sigma\rho^2(y_t^x)\sum_{\delta_1\dots\delta_s} L^3_{\delta_1,\dots,\delta_s}\big\langle\mathcal{D}^{\alpha\cup\beta}y_t^x, \mathcal{D}^{\delta_1}y_t^x\dots\mathcal{D}^{\delta_s}y_t^x\big\rangle\Big|dt$$

$$\leq C\int_0^t \mathbf{E}\,p_i'(\rho^2(y_t^x))\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2q}\sum_{\sigma=1}^d \frac{\|A_\sigma\rho^2(y_t^x)\|^2}{\rho^2(y_t^x)}dt \quad (4.41)$$

$$+C\int_0^t \mathbf{E}p_i'(\rho^2(y_t^x))\rho^2(y_t^x)\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2(q-1)}\|\nabla A_\sigma[\mathcal{D}^{\alpha\cup\beta}y_t^x]\|^2 \qquad (4.42)$$

$$+C\int_0^t \mathbf{E}p_i'(\rho^2(y_t^x))\rho^2(y_t^x)\|\mathcal{D}^{\alpha\cup\beta}y_t^x\|^{2(q-1)}$$

$$\times\sum_{\delta_1\ldots\delta_s} L^3_{\delta_1,\ldots,\delta_s}\|\mathcal{D}^{\delta_1}y_t^x\ldots\mathcal{D}^{\delta_s}y_t^x\|^2 dt. \qquad (4.43)$$

Term (4.41) is added to (4.40) and estimated by $\int_0^t h(\tau)d\tau$ after application of (4.12). The term (4.42) is combined with (4.33), leading, together with (4.32)–(4.34), to the coercitivity condition with constant K_C.

Step 4. To insert weights $1/t^{2q|\beta|}$ into $Q_n(t)$ and return to the proof of nonlinear estimate we need the following generalization of Gronwall-Bellmann lemma, first found in [3, 8].

Lemma 4.1. *Let* $h_0 = 0$ *and*

$$h_t \le M\int_0^t h_t\,dt + \int_0^t [\,L_t + h_t^{1-1/\alpha}K_t\,]\,dt \qquad (4.44)$$

with $L_t, K_t \ge 0$ *such that*

$$\int_0^t \frac{L_s}{s^\alpha}\,ds < \infty \qquad and \qquad \sup_{s\in[0,t]} K_s < \infty$$

. Then

$$\sup_{s\in[0,t]}\frac{h_s}{s^\alpha} \le \alpha e^{Mt}\int_0^t \frac{L_s}{s^\alpha}\,ds + \frac{e^{\alpha Mt}}{\alpha^\alpha}\sup_{s\in[0,t]}K_s^\alpha. \qquad (4.45)$$

First note that due to (4.7) the highest order of behaviour of functions L^7, K^1, L^6 and L^3 in I_1, I_2, I_3 and (4.43) is $(1+\rho^2(y_t^x))^{\mathbf{k}'}$ with some $\mathbf{k}'$, arising from (4.7).

Using inequality

$$\mathbf{E}px^{2q-1}y \le \Big(\mathbf{E}px^{2q}\Big)^{1-1/2q}\Big(\mathbf{E}py^{2q}\Big)^{1/2q}$$

we estimate I_2 by

$$I_2 \le C_1 \int_0^t \sum_{\substack{\beta_1,\dots,\beta_a,\\ \varepsilon_1,\dots,\varepsilon_b}} \Big(\mathbf{E}\, p_i(\rho^2) \|\mathcal{D}^{\alpha\cup\beta} y_t^x\|^{2q}\Big)^{1-1/2q} \Big(\mathbf{E}\, p_i(\rho^2)(1+\rho^2)^{2q\mathbf{k}}$$
$$\times \|\mathcal{D}^{\beta_1} y_t^x \dots \mathcal{D}^{\beta_a} y_t^x \mathcal{D}^{\varepsilon_1\setminus k_1} \widetilde{z}_{k_1} \dots \mathcal{D}^{\varepsilon_b\setminus k_b} \widetilde{z}_{k_b}\|^{2q}\Big)^{1/2q}. \quad (4.46)$$

To obtain form (4.44) we estimate I_1, I_3 and (4.43) using inequality $x^{m-n} y^n \le \frac{m-n}{m} x^m + \frac{n}{m} y^m$ with $m = 2q$ and $n = 1$ or 2. Thus we have

$$h(t) \le C_1 \int_0^t h(\tau)\, d\tau + C_2 \int_0^t L_\tau\, d\tau + C_3 \int_0^t h(\tau)^{1-1/2q} K_\tau\, d\tau, \quad (4.47)$$

where

$$L_\tau = \sum_{\delta_1,\dots,\delta_s} \mathbf{E} p_i (1+\rho^2)^{2q\mathbf{k}} \|\mathcal{D}^{\delta_1} y_t^x \dots \mathcal{D}^{\delta_s} y_t^x\|^{2q}, \quad (4.48)$$

$$K_\tau = \sum_{\substack{\beta_1,\dots,\beta_a,\\ \varepsilon_1,\dots,\varepsilon_b}} \Big(\mathbf{E}\, p_i(\rho^2)(1+\rho^2)^{2q\mathbf{k}}$$
$$\times \Big\|\mathcal{D}^{\beta_1} y_t^x \dots \mathcal{D}^{\beta_a} y_t^x \mathcal{D}^{\varepsilon_1\setminus k_1} \widetilde{z}_{k_1} \dots \mathcal{D}^{\varepsilon_b\setminus k_b} \widetilde{z}_{k_b}\Big\|^{2q}\Big)^{1/2q}. \quad (4.49)$$

Step 5. To apply Lemma 4.1 it remains to estimate L_τ and K_τ terms. To estimate L_τ term we use inequality

$$\|x_{\delta_1} \dots x_{\delta_s}\|^{m/i} \le (\|x_{\delta_1}\|^{m/|\delta_1|})^{|\delta_1|/i} \dots (\|x_{\delta_s}\|^{m/|\delta_s|})^{|\delta_s|/i}.$$

Due to $2q = m/i$ and $\delta_1 \cup \dots \cup \delta_s = \{1,\dots,i\}$ we have

$$\frac{1}{\tau^{2q|\beta|}} L_\tau$$
$$\le \sum_{\delta_1,\dots,\delta_s} \mathbf{E}\, p_i(\rho^2)(1+\rho^2)^{\mathbf{k}m/i} \Big\|\frac{1}{\tau^{|\delta_1'|}} \mathcal{D}^{\delta_1} y_t^x\Big\|^{m/i} \dots \Big\|\frac{1}{\tau^{|\delta_s'|}} \mathcal{D}^{\delta_s} y_t^x\Big\|^{m/i}$$
$$\le \sum_{\delta_1,\dots,\delta_s} \mathbf{E} \Big(p_{|\delta_1|}(\rho^2) \Big\|\frac{1}{\tau^{|\delta_1'|}} \mathcal{D}^{\delta_1} y_t^x\Big\|^{m/|\delta_1|}\Big)^{|\delta_1|/i} \times \dots$$
$$\times \Big(p_{|\delta_s|}(\rho^2) \Big\|\frac{1}{\tau^{|\delta_s'|}} \mathcal{D}^{\delta_s} y_t^x\Big\|^{m/|\delta_s|}\Big)^{|\delta_s|/i}$$
$$\le \sum_{\substack{\delta_1,\dots,\delta_s,\\ s\ge 2}} \mathbf{E} \Big(\frac{|\delta_1|}{i} p_{|\delta_1|}(\rho^2) \Big\|\frac{1}{\tau^{|\delta_1'|}} \mathcal{D}^{\delta_1} y_t^x\Big\|^{m/|\delta_1|} + \dots$$

$$+ \frac{|\delta_s|}{i} p_{|\delta_s|}(\rho^2) \left\| \frac{1}{\tau^{|\delta'_s|}} \mathcal{D}^{\delta_s} y_t^x \right\|^{m/|\delta_s|} \Big), \quad (4.50)$$

where δ' means the stochastic part of set δ. Above we inserted the non-linear hierarchies of polynomial weights (4.9) and used Young inequality $|z_1 \dots z_n| \leq |z_1|^{q_1}/q_1 + \dots + |z_n|^{q_n}/q_n$ with $1/q_1 + \dots + 1/q_n = 1$.

Due to $s \geq 2$ in (4.50) variation $\mathcal{D}^{\alpha \cup \beta} y_t^x$ splits to at least two lower order variations, presented in expression $Q_{n-1}^{n'}$. Therefore

$$\frac{L_\tau}{\tau^{2q|\beta|}} \leq (4.50) \leq C Q_{n-1}^{n'}(y,\tau) + C' Q_n^{n'-1}(y,\tau). \quad (4.51)$$

Step 6. It remains to estimate (4.49) from above. First we estimate K_τ term by

$$K_\tau \leq \sum_{\beta,\varepsilon} \Big(\mathbf{E}\, p_i (1+\rho^2)^{2q\mathbf{k}'} \prod_{j=1}^a \|\mathcal{D}^{\beta_j} y_t^x\|^{2q} \cdot \prod_{j=1}^b \|\mathcal{D}^{\varepsilon_j \backslash k_j} \widetilde{z}_{k_j}\|^{2q} \Big)^{1/2q} \quad (4.52)$$

Consider one term in (4.52) of the type $\|\mathcal{D}^{\varepsilon_j \backslash k_j} \widetilde{z}_{k_j}\|^{2q}$. Due to (3.4)

$$\widetilde{z}_k^\sigma = [A^{-1}(y_t^x)]_p^\sigma \frac{\partial y^p}{\partial x^k}. \quad (4.53)$$

Then

$$\mathcal{D}^{\varepsilon \backslash k} \widetilde{z}_k = \mathcal{D}^{\varepsilon \backslash k} \Big[(A^{-1}) \frac{\partial y}{\partial x^k} \Big]$$
$$= \sum_{\substack{\mu_1 \cup \dots \cup \mu_\ell = \varepsilon \backslash k, \\ |\mu_1| \geq 0}} (A^{-1})^\ell \mathcal{D}^{\mu_1} \frac{\partial y}{\partial x^k} \cdot \mathcal{D}^{\mu_2} y_t^x \dots \mathcal{D}^{\mu_\ell} y_t^x. \quad (4.54)$$

Above summation runs on all subdivisions of set $\varepsilon \backslash k$ on nonintersecting subsets $\mu_1 \cup \dots \cup \mu_\ell = \varepsilon \backslash k$ with $\ell = 1, \dots, |\varepsilon \backslash k|$, and set μ_1 may be empty. Using (4.8) we have

$$\prod_{j=1}^b \|\mathcal{D}^{\varepsilon_j \backslash k_j} \widetilde{z}_{k_j}\|^{2q}$$
$$\leq \prod_{j=1}^b \sum_{\substack{\mu_1 \cup \dots \cup \mu_\ell = \varepsilon_j \backslash k_j, \\ |\mu_1| \geq 0}} (1+\rho^2(y_t^x))^{2q\mathbf{k}_1 \ell} \left\| \mathcal{D}^{\mu_1} \frac{\partial y}{\partial x_j^k} \right\|^{2q} \dots \|\mathcal{D}^{\mu_\ell} y\|^{2q}$$
$$= \sum_{\nu_1 \cup \dots \cup \nu_\ell = \bigcup_{j=1}^b (\varepsilon_j \backslash k_j)} (1+\rho(y_t^x))^{2q\mathbf{k}_1 (\sum_{j=1}^b |\varepsilon_j| - b)} \left\| \mathcal{D}^{\nu_1} \frac{\partial y}{\partial x^{k_1}} \right\|^{2q} \times \dots$$

$$\times \left\| \mathcal{D}^{\nu_b} \frac{\partial y}{\partial x^{k_b}} \right\|^{2q} \|\mathcal{D}^{\nu_{b+1}} y\|^{2q} \ldots \|\mathcal{D}^{\nu_\ell} y\|^{2q}. \quad (4.55)$$

Above we separated first multiplicators $\mathcal{D}^{\mu} \frac{\partial y}{\partial x}$ in (4.54) to the first b terms $\mathcal{D}^{\nu_j} \frac{\partial y}{\partial x^{k_j}}$, $j = 1, \ldots, b$. In the last line of (4.55) index ℓ in the summation runs from 1 to $\sum_{j=1}^{b} |\varepsilon_j \backslash k_j| = \sum_{j=1}^{b} |\varepsilon_j| - b$, and, for the first b terms of the product, the sets $\nu_1, \ldots, \nu_\ell$ may be empty: $|\nu_1|, \ldots, |\nu_b| \geq 0$. For remaining terms $|\nu_{b+1}|, \ldots, |\nu_\ell| \geq 1$.

Using (4.55), we continue estimate (4.52) of K_τ by

$$K_\tau \leq \sum_{\beta_1 \cup \ldots \cup \beta_a \cup \varepsilon_1 \cup \ldots \cup \varepsilon_b = \{1,\ldots,i\}} \Big[\mathbf{E}\, p_i (1 + \rho^2(y_t^x))^{2q\mathbf{k}'} \prod_{j=1}^{a} \|\mathcal{D}^{\beta_j} y\|^{2q}$$

$$\times \sum_{\nu_1 \cup \ldots \cup \nu_\ell = \bigcup_{j=1}^{b} (\varepsilon_j \backslash k_j)} (1+\rho^2(y_t^x))^{2q\mathbf{k}_1(\sum_{j=1}^{b} |\varepsilon_j| - b)} \left\| \mathcal{D}^{\nu_1} \frac{\partial y}{\partial x^{k_1}} \right\|^{2q} \ldots \left\| \mathcal{D}^{\nu_b} \frac{\partial y}{\partial x^{k_b}} \right\|^{2q}$$

$$\times \|\mathcal{D}^{\nu_{b+1}} y\|^{2q} \ldots \|\mathcal{D}^{\nu_\ell} y\|^{2q} \Big]^{1/2q}. \quad (4.56)$$

We choose $\mathbf{k} = \max\Big(\mathbf{k}', \mathbf{k}_1(\sum_{j=1}^{b} |\varepsilon_j| - b)\Big)$ and continue

$$(4.56) = \sum_{\beta_1 \cup \ldots \cup \beta_a \cup \varepsilon_1 \cup \ldots \cup \varepsilon_b = \{1,\ldots,i\}} \Big[\mathbf{E}\, p_i (1 + \rho(y_t^x))^{2q\mathbf{k}}$$

$$\times \sum_{\nu_1 \cup \ldots \cup \nu_\ell = \cup(\varepsilon_j \backslash k_j)} \|\mathcal{D}^{\beta_1} y\|^{2q} \ldots \|\mathcal{D}^{\beta_a} y\|^{2q} \left\| \mathcal{D}^{\nu_1} \frac{\partial y}{\partial x^{k_1}} \right\|^{2q} \times \ldots$$

$$\times \left\| \mathcal{D}^{\nu_b} \frac{\partial y}{\partial x^{k_b}} \right\|^{2q} \cdot \|\mathcal{D}^{\nu_{b+1}} y\|^{2q} \ldots \|\mathcal{D}^{\nu_\ell} y\|^{2q} \Big]^{1/2q}$$

$$\leq C \sum_{\beta_1 \cup \ldots \cup \beta_a \cup \varepsilon_1 \cup \ldots \cup \varepsilon_b = \{1,\ldots,i\}} \; \sum_{\nu_1 \cup \ldots \cup \nu_\ell = \bigcup_{j=1}^{b} (\varepsilon_j \backslash k_j)} \Big[\mathbf{E}\, p_i (1 + \rho^2(y_t^x))^{2q\mathbf{k}}$$

$$\times \|\mathcal{D}^{\beta_1} y\|^{2q} \ldots \|\mathcal{D}^{\beta_a} y\|^{2q} \left\| \mathcal{D}^{\nu_1} \frac{\partial y}{\partial x^{k_1}} \right\|^{2q} \times \ldots$$

$$\times \left\| \mathcal{D}^{\nu_b} \frac{\partial y}{\partial x^{k_b}} \right\|^{2q} \|\mathcal{D}^{\nu_{b+1}} y\|^{2q} \ldots \|\mathcal{D}^{\nu_\ell} y\|^{2q} \Big]^{1/2q}. \quad (4.57)$$

Now, due to $\bigcup_{j=1}^{b} (\varepsilon_j \backslash k_j) = \{1, \ldots, i\} \backslash \big(\beta_1 \cup \ldots \cup \beta_a \cup \{k_1, \ldots, k_b\}\big)$, we rewrite the double sum as

$$\sum_{\beta_1\cup\ldots\cup\beta_a\cup\varepsilon_1\cup\ldots\varepsilon_b=\{1,\ldots,i\}} \sum_{\nu_1\cup\ldots\cup\nu_\ell=\bigcup_{j=1}^{b}(\varepsilon_j\setminus k_j)} = \sum_{\beta_1\cup\ldots\cup\beta_a\cup\nu_1\cup\ldots\cup\nu_\ell=\{1,\ldots,i\}\setminus\{k_1,\ldots,k_b\}} = \sum_{\mu_1\cup\ldots\cup\mu_\ell=\{1,\ldots,i\}\setminus\{k_1,\ldots,k_b\}}.$$

The last sum runs on all subdivisions of set $\{1,\ldots,i\}\setminus\{k_1,\ldots,k_b\}$ on nonintersecting subsets $\mu_1,\ldots,\mu_\ell$ such that $|\mu_1|,\ldots,|\mu_b|\geq 0$ and $|\mu_{b+1}|,\ldots,|\mu_\ell|\geq 1$.

Thus we have

$$(4.57)\leq \sum_{\mu_1\cup\ldots\cup\mu_\ell=\{1,\ldots,i\}\setminus\{k_1,\ldots,k_b\}} \Big[\mathbf{E}\, p_i(1+\rho^2(y_t^x))^{2q\mathbf{k}} \times \Big\|\mathcal{D}^{\mu_1}\frac{\partial y}{\partial x^{k_1}}\Big\|^{2q}\ldots\Big\|\mathcal{D}^{\mu_b}\frac{\partial y}{\partial x^{k_b}}\Big\|^{2q}\cdot\|\mathcal{D}^{\mu_{b+1}}y\|^{2q}\ldots\|\mathcal{D}^{\mu_\ell}y\|^{2q}\Big]^{1/2q}. \tag{4.58}$$

We see that the general number of stochastic derivatives in (4.58) is reduced on b (because due to (4.53) the stochastic indexes k in $\widetilde{z}_k$ were replaced by ordinary variations), in comparison to the initial nonlinear expression $Q_n^{n'}$. So in this terms the stochastic indexes are transformed to the ordinary variations and order of ordinary derivatives increases at least at 1 in comparison to the initial nonlinear expression $Q_n^{n'}$. Therefore, proceeding like in (4.50) with the use of hierarchy of polynomial weights (4.9) and property $|\mu_1|+\ldots+|\mu_\ell|=i-b$, we come to the lower stochastic order nonlinear expression

$$\begin{aligned}
(4.58)\leq &\sum_{\mu_1\cup\ldots\cup\mu_\ell=\{1,\ldots,i\}\setminus\{k_1,\ldots,k_b\}} \Big[\mathbf{E}\, p_i(1+\rho^2(y_t^x))^{2q\mathbf{k}} \\
&\qquad\times\Big\|\frac{1}{\tau^{|\mu_1'|}}\mathcal{D}^{\mu_1}\frac{\partial y}{\partial x^{k_1}}\Big\|^{2q}\ldots\Big\|\frac{1}{\tau^{|\mu_\ell'|}}\mathcal{D}^{\mu_\ell}y\Big\|^{2q}\Big]^{1/2q} \\
\leq e^{Mt}&\sum_{\mu_1,\ldots,\mu_\ell}\Big[\Big(\mathbf{E}\, p_{|\mu_1|+1}\Big\|\frac{1}{\tau^{|\mu_1'|}}\mathcal{D}^{\mu_1}\frac{\partial y}{\partial x^{k_1}}\Big\|^{m/(|\mu_1|+1)}\Big)^{(|\mu_1|+1)/i}\times\ldots \\
&\qquad\times\Big(\mathbf{E}\, p_{|\mu_b|+1}\Big\|\frac{1}{\tau^{|\mu_b'|}}\mathcal{D}^{\mu_b}\frac{\partial y}{\partial x^{k_b}}\Big\|^{m/(|\mu_b|+1)}\Big)^{(|\mu_b|+1)/i} \\
&\qquad\Big(\mathbf{E}\, p_{|\mu_{b+1}|}\Big\|\frac{1}{\tau^{|\mu_{b+1}'|}}\mathcal{D}^{\mu_{b+1}}y\Big\|^{m/|\mu_{b+1}|}\Big)^{|\mu_{b+1}|/i}\times\ldots \\
&\qquad\times\Big(\mathbf{E}\, p_{|\mu_\ell|}\Big\|\frac{1}{\tau^{|\mu_\ell'|}}\mathcal{D}^{\mu_\ell}y\Big\|^{m/|\mu_\ell|}\Big)^{|\mu_\ell|/i}\Big]^{1/2q} \\
\leq e^{Mt}&\sum_{\mu_1,\ldots,\mu_\ell}\Big[\mathbf{E}\Big(\frac{|\mu_1|+1}{i}p_{|\mu_1|+1}\Big\|\frac{1}{\tau^{|\mu_1'|}}\mathcal{D}^{\mu_1}\frac{\partial y}{\partial x^{k_1}}\Big\|^{m/(|\mu_1|+1)}+\ldots
\end{aligned}$$

$$+\frac{|\mu_\ell|}{i}p_{|\mu_\ell|}\Big\|\frac{1}{\tau^{|\mu'_\ell|}}\mathcal{D}^{\mu_\ell}y\Big\|^{m/|\mu_\ell|}\Big)\Big]^{1/2q} \le Ce^{M\,t}\Big(\sum_{\pi=1}^{b}Q_{n+\pi}^{n'-\pi}(y,\tau)\Big)^{1/2q}. \tag{4.59}$$

Here we multiplied and divided on τ^{i-b} and applied estimate $\tau^{i-b}\le e^{(i-b)\tau}$.

Step 7. Summarizing steps 5–6, we have estimate on one of terms in expression $Q_n^{n'}$ (4.3):

$$\sup_{s\in[0,t]}\mathbf{E}\,p_i\Big\|\frac{1}{s^{|\beta|}}\mathcal{D}^{\alpha\cup\beta}y_s^x\Big\|^{m/i}$$
$$\le Ce^{(M+1)t}\sup_{s\in[0,t]}(Q_{n-1}^{n'}(s)+Q_n^{n'-1}(s)+\sum_{\pi=1}^{b}Q_{n+\pi}^{n'-\pi}(s)),$$

where we used estimates (4.51), (4.59) on terms L_τ, K_τ in (4.45). Therefore

$$\sup_{s\in[0,t]}Q_n^{n'}(s)\le C'e^{M't}\sup_{s\in[0,t]}(Q_{n-1}^{n'}(s)+Q_n^{n'-1}(s)+\sum_{\pi=1}^{b}Q_{n+\pi}^{n'-\pi}(s)).$$

Iterating the above recurrence, one comes to

$$\sup_{s\in[0,t]}Q_n^{n'}(s)\le Ke^{M\,t}\sup_{s\in[0,t]}Q_{n+n'}^{0}(s).$$

The last expression $Q_{n+n'}^0$ does not contain stochastic derivatives, therefore we can apply the nonlinear estimate on ordinary variations $Q_m(s)\le e^{M\,s}Q_m(0)$, found in [6]. The last estimate could be also seen in a lines of current proof, if one neglects all complications with stochastic derivatives and uses standard Gronwall-Bellmann inequality after (4.43). □

5. Regular properties of semigroup.

Now, by application of nonlinear estimates we get the smooth and raise of smoothness estimates on semigroup. Introduce the space of continuously differentiable functions $C^n_{\vec{q}(\mathbf{k})}(M)$, equipped with norm

$$\|f\|_{C^n_{\vec{q}(\mathbf{k})}}=\max_{i=0,\dots,n}\sup_{x\in M}\frac{\|(\nabla^x)^i f(x)\|}{q_i(\rho^2(x,o))}.$$

Monotone strictly positive weights q_i of polynomial behaviour fulfill hierarchy

$$\exists\, \mathbf{k}\ \forall\, i = 1, \ldots, n\ \forall\, u \geq 0 \quad q_{i+1}(u) \geq (1 + |u|)^{\mathbf{k}} q_i(u). \tag{5.1}$$

Due to the triangle inequality for metric the choice of particular point $o \in M$ becomes inessential.

Main statement is the raise of smoothness property of semigroup in scale $C^n_{\vec{q}(\mathbf{k})}$.

Theorem 5.1. *Suppose conditions* (4.4)–(4.8) *hold. Then there is* $\widetilde{\mathbf{k}} \geq \mathbf{k}$ *and* $K,\ N(n, m, \mathbf{k})$ *such that*

$$\forall\, t > 0 \quad P_t : C^n_{\vec{q}(\mathbf{k})}(M) \to C^{n+m}_{\vec{q}(\mathbf{k}), q_{n+1}(\mathbf{k}), \ldots, q_{n+m}(\mathbf{k})}(M)$$

and smooth $(m = 0)$ *and raise of smoothness* $(m > 0)$ *estimates hold*

$$\|P_t f\|_{C^{n+m}_{\vec{q}(\mathbf{k}), q_{n+1}(\mathbf{k}), \ldots, q_{n+m}(\mathbf{k})}} \leq \frac{K e^{Nt}}{t^{m/2}} \|f\|_{C^n_{\vec{q}(\mathbf{k})}}.$$

Proof. First we discuss the raise of smoothness estimates. Remark that the statement of theorem in the raise of smoothness part is recursive on m. Indeed, by multiplicative property of semigroup $P_t = (P_{t/m})^m$ and inequality

$$\begin{aligned}\|P_t f\|_{C^{n+m}} &= \|(P_{t/m})^m f\|_{C^{n+m}} \leq \frac{K' e^{N' t}}{\sqrt{t}} \|(P_{t/m})^{m-1} f\|_{C^{n+m-1}} \\ &\leq \frac{K' K'' e^{(N'+N'')t}}{t} \|(P_{t/m})^{m-2} f\|_{C^{n+m-2}} \leq \ldots \leq \frac{K e^{Nt}}{t^{m/2}} \|f\|_{C^n}.\end{aligned}$$

Therefore it is sufficient to demonstrate the first order raise of smoothness estimate.

Raise of smoothness representation. To show that semigroup acts from C^i to C^{i+1} we take one of representations (3.1) and differentiate it one time.

$$\begin{aligned}\nabla^x (\nabla^x)^i P_t f(x) &= \sum_{\substack{j_1 + \ldots + j_s = i+1, \\ s \geq 1}} \mathbf{E} \langle (\nabla^y)^s f(y^x_t), (\mathbb{V}^x)^{j_1} y^x_t \otimes \ldots \otimes (\mathbb{V}^x)^{j_s} y^x_t \rangle \\ &= \mathbf{E} \langle (\nabla^y)^{i+1} f(y^x_t), \mathbb{V}^x y^x_t \otimes \ldots \otimes \mathbb{V}^x y^x_t \rangle + \{\text{terms with } s \geq 2\}. \end{aligned} \tag{5.2}$$

Then we use that by (3.3)

$$(\nabla^y)^{i+1} f(y^x_t) \mathbb{V}^x y^x_t = \mathbb{V}^x (\nabla^y)^i f(y^x_t) = \frac{1}{t} \widetilde{I\!\!D} (\nabla^y)^i f(y^x_t)$$

and, similar to (3.4)-(3.8), disintegrate by parts (3.6) terms with the highest order $(i+1)^{th}$ derivative of initial function

$$\nabla^x(\nabla^x)^i P_t f(x) = \mathbf{E}\Big\langle (\nabla^y)^i f(y_t^x), \Big(\frac{1}{t}\widetilde{I\!\!D} + \frac{1}{t}\int_0^t \widetilde{z}^\sigma dW^\sigma\Big)(\mathbb{V}^x y_t^x \otimes \ldots \otimes \mathbb{V}^x y_t^x)\Big\rangle \tag{5.3}$$

$$+ \sum_{j_1+\ldots+j_s=i+1,\ s\geq 2} \mathbf{E}\langle (\nabla^y)^s f(y_t^x), (\mathbb{V}^x)^{j_1} y_t^x \otimes \ldots \otimes (\mathbb{V}^x)^{j_s} y_t^x\rangle. \tag{5.4}$$

Choice of weights p_i and unification of estimation on (5.4)–(5.3) terms. An easy check demonstrates that the choice of weights $p_j(u) = P(u)(1+|u|)^{m\mathbf{k}(1/j-1/i)}$ fulfills hierarchy (4.9). For this choice $\widetilde{p_i} = P$, therefore we have from nonlinear estimate (4.10) that

$$\begin{aligned}\mathbf{E}P(\rho^2(y_t^x, o))\,\|\frac{1}{t^{|\beta|}}(I\!\!D)^i y_t^x\|^{q/i} &\leq Ke^{Nt}Q_i^0(y,0)\\ &= Ke^{Nt}P(\rho^2(x,o))(1+\rho^2(x,o))^{\mathbf{k}q(i-1)/i},\end{aligned} \tag{5.5}$$

where, in order to find $Q_i^0(y,0)$, we used that the initial data for ordinary variations are

$$\mathbb{V}_k(y_t^x)^m\Big|_{t=0} = \frac{\partial x^m}{\partial x^k} = \delta_k^m,$$
$$\mathbb{V}_{kj}(y_t^x)^m\Big|_{t=0} = \mathbb{V}_k(\mathbb{V}_j y^m)\Big|_{t=0} = \partial_k(\delta_j^m) - \Gamma_{k\ j}^{\ h}(x)\delta_h^m + \Gamma_{p\ q}^{\ m}(y_0^x)\delta_j^p\delta_k^q = 0,$$
$$(\mathbb{V}^x)^i y_t^x\Big|_{t=0} = 0, \quad \forall\, i\geq 1.$$

Next we remark that all terms in (5.4)–(5.3) have form

$$\mathbf{E}\langle (\nabla^y)^s f(y_t^x), \eta^{(j_1)} \otimes (\mathbb{V}^x)^{j_2} y_t^x \otimes \ldots \otimes (\mathbb{V}^x)^{j_s} y_t^x\rangle, \tag{5.6}$$

where $\eta^{(j_1)}$ represents

1. High order variation $(\mathbb{V}^x)^{j_1} y_t^x$ for (5.4);
2. Stochastic variation $\frac{1}{t}\widetilde{I\!\!D} y_t^x$ for $(5.3)_1$;
3. Stochastic integral $\int_0^t \widetilde{z}^\sigma dW^\sigma$ for $(5.3)_2$.

Raise of smoothness estimates. Now we can apply (5.5) to estimate derivatives (5.6) in topologies $C^n_{\vec{q}}(M)$

$$\frac{\|(\nabla^x)^{i+1}P_tf(x)\|_{T_x^{(0,i+1)}}}{q_{i+1}(\rho^2(x,o))} \leq$$

$$\sum_{terms\ (5.6)} \frac{\|\mathbf{E}\left\langle (\nabla^y)^s f(y_t^x)\,, \eta^{(j_1)} \otimes (\mathbb{\nabla}^x)^{j_2} y_t^x \otimes \ldots \otimes (\mathbb{\nabla}^x)^{j_s} y_t^x \right\rangle\|}{q_{i+1}(\rho^2(x,o))}$$

$$\leq \sum_{terms} \left(\sup_{y_t^x \in M} \frac{\|(\nabla^y)^s f(y_t^x)\|_{T_y^{(0,s)}}}{q_s(\rho^2(y_s^x,o))} \right)$$

$$\times \frac{\mathbf{E} q_s(\rho^2(y_s^x,o)) \|\eta^{(j_1)}\| \, \|(\mathbb{\nabla}^x)^{j_2} y_t^x\| \ldots \|(\mathbb{\nabla}^x)^{j_s} y_t^x\|}{q_{i+1}(\rho^2(x,o))}$$

$$\leq \sum_{terms} \|f\|_{C^n_{\vec{q}}} \frac{\left(\mathbf{E} q_s(\rho^2(y_s^x,o)) \|\eta^{(j_1)}\|^{i/j_1}\right)^{j_1/i}}{q_{i+1}(\rho^2(x,o))}$$

$$\times \prod_{\ell=2}^{s} \left(\mathbf{E} q_s(\rho^2(y_s^x,o)) \|(\mathbb{\nabla}^x)^{j_\ell} y_t^x\|^{i/j_\ell}\right)^{j_\ell/i}, \quad (5.7)$$

where we substituted intermediate weights $q_s(\rho^2)$ and at last step applied Holder inequality.

Because term with η is ordinary or stochastic variation with factor $1/t$, the nonlinear estimate in form (5.5) applies. The estimation of the stochastic integral reduces to the nonlinear estimate on the first order variation, if one uses standard estimate

$$\mathbf{E}\left(\int_0^t \widetilde{z}_s^\sigma \, dW_s^\sigma \right)^{2q} \leq K_q t^{q-1} \mathbf{E} \int_0^t \|\widetilde{z}_s\|^{2q} ds \quad (5.8)$$

and recalls representation (3.4) and condition (4.8)

$$(5.8) \leq K_q t^{q-1} \mathbf{E} \int_0^t (1+\rho^2(y_s^x,o))^{2q\mathbf{k}_1} \, \|\mathbb{\nabla}^x y_t^x\|^{2q} ds$$

$$\leq K\, t^q e^{Nt} (1+\rho^2(x,o))^{2q\mathbf{k}_1}.$$

Above we applied again applied (5.5).

Therefore the last fraction in (5.7) is estimated by (5.5)

$$\frac{(\mathbf{E} q_s(\rho^2(y_s^x,o)) \|\eta^{(j_1)}\|^{i/j_1})^{j_1/i}}{q_{i+1}(\rho^2(x,o))} \prod_{\ell=2}^{s} (\mathbf{E} q_i(\rho^2(y_s^x,o)) \|(\mathbb{\nabla}^x)^{j_\ell} y_t^x\|^{i/j_\ell})^{j_\ell/i}$$

$$\leq \frac{1}{q_{i+1}(\rho^2(x,o))} \cdot \left\{ \text{factor} \quad \begin{matrix} 1 \text{ for } \eta \text{ of types 1,2} \\ \dfrac{1}{\sqrt{t}} \text{ for } \eta \text{ of type 3} \end{matrix} \right\}$$

$$\times\,(e^{Nt}q_s(\rho^2(x,o))(1+\rho^2(x,o))^{\mathbf{k}_1 i/j_1+\mathbf{k}i(j_1-1)/j_1})^{j_1/i}$$

$$\times\prod_{\ell=2}^{s}(e^{Nt}q_s(\rho^2(x,o))(1+\rho^2(x,o))^{\mathbf{k}i(j_\ell-1)/j_\ell})^{j_\ell/i}$$

$$=\frac{\mathrm{const}}{\sqrt{t}}e^{Nt}\frac{q_s(\rho^2(x,o))(1+\rho^2(x,o))^{\mathbf{k}_1+\mathbf{k}(i+1-s)}}{q_{i+1}(\rho^2(x,o))}\le\frac{K\,e^{Nt}}{\sqrt{t}},$$

where we used hierarchy (5.1) with additional weight $\widetilde{\mathbf{k}}=\mathbf{k}+\mathbf{k}_1$ for the differentiability order $n+1$. Moreover, applying (5.5), we used that in notations of (4.3) $q=i,\ n=j_\ell$ and $j_1+\ldots+j_s=i+1$.

This leads to the statement.

2. The smooth estimates ($m=0$) are demonstrated in [6]. They are verified like above from representations (3.1), without additional $(i+1)^{th}$ differentiation and integration by parts like in Step 1.

The final conclusion about continuous differentiability of semigroup follows from estimates on the continuity in mean of variational processes with respect to the initial data. This fact can be proved under conditions (4.4)–(4.8) in a similar to nonlinear estimate (4.10) way with application of symmetries (4.2) to deal with non-Lipschitz coefficients, like in e.g. [2, 3].

This becomes possible, because, by Theorem 2.1 and asymptotics (4.7), variational equations represent non-autonomous and inhomogeneous equation with respect to the high order variation, if all lower order variations are already constructed. The behaviour of non-autonomous part is controlled by dissipativity and coercitivity condition. In a similar way the nonlinear symmetries (4.2) and polynomial behaviour of coefficients (4.7) give a set of optimal estimates on inhomogeneous part, like in (4.10). Therefore, like in [2, 3], variational processes are easy constructed as strong solutions to systems (2.5), (2.7)–(2.8)

Turning to the C^∞ differentiability of process y_t^x on initial data, it is necessary to demonstrate that the solutions of variational equations represent high order invariant derivatives of process y_t^x. By schemes of [2, 3] this can be obtained by application of nonlinear symmetries (4.2) in a recurrent on the order of differentiation way. However, because we work in the finite-dimensional situation, we can also apply more stochastic in nature techniques of stopping times, e. g. [17], that guarantee that derivatives of finite dimensional process with locally C^∞ coefficients are represented as solutions to corresponding variational equations before exit times. □

References

[1] A. Val. Antoniouk, A. Vict. Antoniouk, *How the unbounded drift shapes the Dirichlet semigroups behaviour of non-Gaussian Gibbs measures* // J. Funct. Anal. **135** (1996), 488–518.

[2] A. Val. Antoniouk, A. Vict. Antoniouk, *Nonlinear calculus of variations and regularity of parabolic problems*, to appear in collection of selected papers of the 1^{st} Ukrainian Congress of Mathematicians.

[3] A. Val. Antoniouk, A. Vict. Antoniouk, *Nonlinear estimates approach to the regularity properties of diffusion semigroups*, In "Nonlinear Analysis and Applications: To V. Lakshmikantham on his 80th birthday", in two volumes, Eds. Ravi P. Agarwal and Donal O'Regan, Kluwer, **1** (2003), 165–226.

[4] A. Val. Antoniouk, *Nonlinear Symmetries of Variational Calculus and Regularity Properties of Differential Flows on Non-Compact Manifolds*, to appear in Proceedings of 5^{th} Intern. Conf. "Symmetry in nonlinear mathematical physics", 2003, 9 pp.

[5] A. Val. Antoniouk, *Nonlinear Estimates on Regularity of Differential Flows on Manifolds*, to appear in Mathematical methods and physical-mechanical fields, 8 pp.

[6] A. Val. Antoniouk, *Regularity of nonlinear flows on manifolds: nonlinear estimates on variations or why generalizations of classical covariant derivatives are required?*, to appear in Nonlinear Boundary Problems, 2004, 12 pp.

[7] A. Val. Antoniouk, A. Vict. Antoniouk, *High order formulae for derivatives of nonlinear diffusion semigroups*, Ukrainian Math. Journal, January 2001, No 1, 117–122.

[8] A. Val. Antoniouk, A. Vict. Antoniouk, *Nonlinear Estimates on Regularity.* Collection of papers, Kiev Inst. Math., 2000.

[9] A. Val. Antoniouk, *Towards upper bounds on general second order operators on metric function*, to appear in Ukrainian Math. Bulletin, 8 pp.

[10] Y. I. Belopolskaja, Y. L. Daletskii, *Stochastic equations and differential geometry.* Kluwer, 1996.

[11] J.-M. Bismut, *Large deviations and the Malliavin calculus.* Progress in Mathematics, v. 45, Birkhäuser, 1984.

[12] K. D. Elworthy, *Stochastic flows on Riemannian manifolds*, in Diffusion Processes and Related Problems in Analysis (M. A. Pinsky and V. Wihstutz, Eds.), Birkhäuser, Basel, **II** (1992), 37–72.

[13] M. Emery, *Stochastic calculus in manifolds.* Springer-Verlag, 1989.

[14] Yu. Gliklikh, *Global analysis in mathematical physics. Geometric and stochastic methods.* Kluwer, 1996.

[15] K. Itô, *Stochastic differential equations in a differentiable manifold* // Nagoya Math. J., **1** (1950), 35–47.

[16] N. V. Krylov, B. L. Rozovskii, *On the evolutionary stochastic equations* // Ser. "Contemporary problems of Mathematics", VINITI, Moscow, **14** (1979), 71–146.

[17] H. Kunita, *Stochastic flows and stochastic differential equations.* Cambridge Uni. Pess, 1990.

[18] P. Malliavin, *Stochastic Analysis*, Springer Verlag, 1997.

[19] E. Pardoux, *Stochastic partial differential equations and filtering of diffusion processes* // Stochastics, **3** (1979), 127–167.

[20] J. F. Pommare, *Systems of partial differential equations and Lie pseudogroups.* Gordon and Breach Sci. Publ., 1978.

[21] L. C. G. Rogers, D. Williams, *Diffusions, Markov processes and martingales*, v. 1, 2, Wiley: New York, 1987.

[22] L. Schwartz, *Semimartingales and their stochastic calculus on manifolds.* Press de l'Universite de Montreal, 1984.

[23] D. W. Stroock, *An introduction to the analysis of paths on Riemannian manifold.* Math. Surveys and Monographs, v. 74, AMS, 2000.

Contact information

Alexander Val. Antoniouk
Alexandra Vict. Antoniouk

Department of Nonlinear Analysis,
Institute of Mathematics NAS of Ukraine,
Tereschenkivska 3,
01601 MSP Kiev-4,
Ukraine
E-Mail: antoniouk@imath.kiev.ua