ON DETERMINISTIC AND STOCHASTIC SLIDING MODES VIA SMALL DIFFUSION APPROXIMATION

GRIGORI N. MILSTEIN, ALEXANDRE YU. VERETENNIKOV

Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstr. 39, D-10117 Berlin , Germany
e-mail: milstein@wias-berlin.de

Institute of Information Transmission Problems
19 Bolshoy Karetnii, 101447 Moscow, Russia
e-mail: ayu@sci.lpi.ac.ru

1991 *Mathematics Subject Classification.* 34A60, 60H10.

Key words and phrases. differential equations with discontinuous right-hand sides, sliding modes, stochastic differential equations, stochastic averaging principle.

[1]This work for the second author was supported by RFBR grant 98-01-0062 and by DFG grant 436 Rus 17/133/97

Abstract. We study solutions of a system of ordinary differential equations with discontinuity of its vector field on a smooth surface via small additive diffusion perturbations. When a diffusion term tends to zero, one obtains limiting sliding modes on the surface with explicit representation for its motion law. Stochastic sliding modes are also established.

1991 Mathematics Subject Classification. 34A60, 60H10.

Key words and phrases. Differential equations with discontinuous right-hand sides, sliding modes, stochastic differential equations, stochastic averaging principle, stochastic differential equations with small noise.

1. **Introduction**

The goal of the paper is to expose an approach to sliding modes via approximation by small diffusion. Let us consider an autonomous d-dimensional system $(d \geq 2)$

$$\frac{dx}{dt} = f(x) \tag{1.1}$$

with the velocity vector field $f(x)$ which is discontinuous on some $(d-1)$-dimensional surface S. Even the definition of solutions of such an equation requires a special examination. Indeed, one should introduce the rule of constructing the trajectory on the surface using additional conventions, so to say, axiomatically, because they do not follow from the original equation (1.1). Various approaches and models were considered in [Filippov (1985), see references therein as well]. We propose one more model which appears to be natural from the "stochastic" point of view. We suggest to think of the equation (1.1) as a limiting one, obtained from a more complicated equation with an additional small "white noise". Obviously, white noise is not a unique possibility and one can try other versions. In some sense, the choice of the approximation model is up to the situation, or physics of the matter. In various problems an approach based on a small diffusion noise is more or less standard. Now we try to apply it to the sliding mode problem.

Assume that the surface S is smooth and any solution of equation (1.1) approaches it from both its sides for a finite time, at least, in some neighborhood of the surface.

We propose the approximative equation

$$dX_t = f(X_t)dt + \varepsilon c(X_t)dw(t) \tag{1.2}$$

with nondegenerate diffusion εc. Here $w(t)$ is a d-dimensional standard Wiener process, ε is a small parameter and c is a nondegenerate $d \times d$-matrix. In fact, in the most simple case c is a constant matrix. We shall start with this case. However, the reader will see that soon one should use the change of variables. Hence, it is reasonable to introduce a general c at the very beginning. Now we are going to study the limit behavior of solutions as $\varepsilon \to 0$. Explicit formulas for limiting coefficients of the system on the surface S will be established.

Obviously, we have in mind that equation (1.2) with bounded f and "good" c has a unique strong solution for any fixed initial data [Veretennikov (1980)]. We always assume c nondegenerate. On the other hand, it can be satisfactory as well to have a weak solution which is unique in law [Krylov (1969), Stroock-Varadhan (1979)]. One could also consider even arbitrary (nondegenerate) discontinuous c which still allows one to have a weak solution which is a homogeneous strong Markov process [Krylov (1973)]. If the limiting behavior is well-defined, in fact, one cannot worry too much even about strong or weak uniqueness before passing to the limit.

Main results are contained in Sections 2 and 4. In Section 2 we study deterministic sliding modes and in Section 4 stochastic sliding modes. A stochastic sliding mode arises if the original system

$$dX_t = f(X_t)dt + \sigma(X_t)dW(t) \tag{1.3}$$

is stochastic and such that its coefficients are discontinuous on S. In addition, its solutions approach S for a finite time in a certain sense and then cannot leave S. Thus, we have two media separated by S, with two different rules of stochastic motions in those media. The stochastic sliding mode problem is again to obtain a "natural" law of a motion on S. Below we present both deterministic and stochastic sliding modes on a $(d-1)$-dimensional surface S of discontinuity in an explicit form. Notice that in the case of deterministic sliding mode our answer coincides with the well-known rule [1]. On the contrary, the stochastic sliding modes seem to be considered for the first time in this paper.

The approach uses essentially the stochastic averaging principle. We first demonstrate this principle for appropriate model problems in details and then state results for more general situations.

In Section 3 we find a sliding mode for a system whose all trajectories approach S for a finite time as above; however, now they attain S with either zero or infinite normal component of velocity (unlike to Section 2).

In Section 5 we consider a model problem of determining a sliding mode on a $(d-2)$-dimensional surface which is the intersection of two $(d-1)$-dimensional sliding surfaces.

We stress out that our exposition is not absolutely rigorous in all parts. In some proofs we give, in fact, a sketch and main ideas. In each case a strict proof is clear. However, it requires much more details and additional pages. The model problems are strict, at any rate.

2. Deterministic sliding mode on $(d-1)$-dimensional surfaces of discontinuity

Let us start with a two-dimensional autonomous system

$$\dot{x} = a(x,y), \ \dot{y} = b(x,y), \tag{2.1}$$

where the functions $a(x,y)$ and $b(x,y)$ are discontinuous on a sufficiently smooth curve S given by an equation $\varphi(x,y) = 0$.

The curve S bisects its neighborhood in the space (x,y) into two domains G^- and G^+. Suppose the right-hand sides of the system (2.1) together with their first derivatives are continuous functions in G^- and G^+ up to S.

Let $(x,y) \in S, \ N = \nabla\varphi(x,y)$ be the normal to S at the point (x,y). Without loss of generality one may suggest N directed to G^+. Let $a^-(x,y), \ b^-(x,y), \ (x,y) \in S$, be the limit values of $a(x,y), \ b(x,y)$ in G^- and $a^+(x,y), \ b^+(x,y)$ be the limit values in G^+. We assume that the projections of the vectors (a^-, b^-) and (a^+, b^+) to the normal N are positive and negative correspondingly, i.e., all solutions of the system (2.1) close to the curve S approach it from both sides with the growth of time and no solution can leave the curve S.

Fix $(x_0, y_0) \in S$. Suppose that the curve S in a neighborhood of (x_0, y_0) can be expressed by the equation $y = \psi(x)$ and the points (x,y) such that $y > \psi(x)$ $(y < \psi(x))$ belong to G^+ (G^-).

Consider the change of variables

$$X = x - x_0,\ Y = y - \psi(x).$$

We get

$$\frac{dX}{dt} = a(X + x_0, Y + \psi(X + x_0)) := \tilde{a}(X, Y)$$

$$\frac{dY}{dt} = b(X + x_0, Y + \psi(X + x_0))$$

$$-\psi'(X + x_0) \cdot a(X + x_0, Y + \psi(X + x_0)) := \tilde{b}(X, Y). \tag{2.2}$$

Then the equation of S in a neighborhood of the origin is $Y = 0$, the domain G^+ is the upper halfplane $Y > 0$ and the domain G^- is the lower halfplane $Y < 0$, $\tilde{b}^-(0,0) = b^-(x_0, y_0) - \psi'(x_0) \cdot a^-(x_0, y_0) > 0$, $\tilde{b}^+(0,0) = b^+(x_0, y_0) - \psi'(x_0) \cdot a^+(x_0, y_0) < 0$.

2.1. Model problem ($d = 2$). In fact, the problem of determining the motion on S has a local nature. Hence, we replace the functions $\tilde{a}$ and $\tilde{b}$ in a small neighborhood of the origin by constant ones, $a^\pm := \tilde{a}^\pm(0,0)$, $b^\pm := \tilde{b}^\pm(0,0)$ and consider the simplified system

$$\dot{X} = a(Y),\ \dot{Y} = b(Y), \tag{2.3}$$

where

$$a(Y) = a(sign\ Y) = \begin{cases} a^-,\ Y < 0, \\ 0,\ Y = 0, \\ a^+,\ Y > 0, \end{cases},\ b(Y) = b(sign\ Y) = \begin{cases} b^-,\ Y < 0, \\ 0,\ Y = 0, \\ b^+,\ Y > 0, \end{cases} \tag{2.4}$$

with $b^- > 0$, $b^+ < 0$.

The method of small additive diffusion suggests that one should consider the following perturbed system of stochastic differential equations:

$$dX_t^\varepsilon = a(Y_t^\varepsilon)dt + \varepsilon c_1 dw_1(t)$$

$$dY_t^\varepsilon = b(Y_t^\varepsilon)dt + \varepsilon c_2 dw_2(t), \tag{2.5}$$

where $c_1$and c_2 are some numbers and also $c_2 \neq 0$, w_1 and w_2 are one-dimensional standard Wiener processes (they are not necessarily independent).

Now we are going to study the limit of the solution with the initial data $(0,0)$ as ε tends to zero.

Theorem 2.1. *Let $(X_t^\varepsilon, Y_t^\varepsilon)$ be the solution of the system (2.5) starting from the origin. Then for any $t > 0$ there exists a limit a.s., in L_1, and in probability in $C([0,T]; R^1)$ for any $T > 0$*

$$\lim_{\varepsilon \to 0} X_t^\varepsilon = (a^- p^- + a^+ p^+) \cdot t := \bar{a} \cdot t, \tag{2.6}$$

where

$$p^- = \frac{b^+}{b^+ - b^-}\ ,\ p^+ = \frac{b^-}{b^- - b^+}\ . \tag{2.7}$$

The probabilities p^- and p^+ are equal to $\mu(Y < 0)$ and $\mu(Y > 0)$ correspondingly where μ is the invariant measure of the Markov process governed by the stochastic differential equation

$$dY = b(Y)ds + c_2 dw(s). \tag{2.8}$$

The measure μ has the following density

$$p(y) = \begin{cases} C \exp \dfrac{2b^- y}{c_2^2}, \ y < 0, \\ \\ C \exp \dfrac{2b^+ y}{c_2^2}, \ y > 0, \end{cases} \qquad C = \frac{2b^- \cdot b^+}{c_2^2(b^+ - b^-)} \ . \tag{2.9}$$

Proof. Evidently, the family of distributions of processes $(X_t^\varepsilon, \ 0 \le t \le T, \ \varepsilon \le 1)$ is relatively compact for any $T > 0$. We shall show now that there is only one limiting point of this family as $\varepsilon \to 0$.

Let us introduce the process

$$Y_s = \varepsilon^{-2} Y^\varepsilon_{\varepsilon^2 s} \ . \tag{2.10}$$

Due to the equality $b(y) = b(sign \ y)$, the law of the process Y_s does not depend on ε and it satisfies the equation (2.8) with a new Wiener process $w(s) = \varepsilon^{-1} w_2(\varepsilon^2 s)$. So we can omit the index ε without a risk of confusion. The Markov process defined by the stochastic differential equation (2.8) is ergodic (cf., [7]). Let μ be its invariant measure.

From the first equation of the system (2.5) we find due to the ergodicity (remind that $X_0^\varepsilon = 0$):

$$X_t^\varepsilon = \int_0^t a(Y_t^\varepsilon) dt + \varepsilon c_1 w_1(t)$$

$$= \int_0^t a(\varepsilon^2 Y_{t/\varepsilon^2}) dt + \varepsilon w_1(t) = \int_0^t a(sign \ Y_{t/\varepsilon^2}) dt + \varepsilon c_1 w_1(t)$$

$$= \frac{\varepsilon^2}{t} \int_0^{t/\varepsilon^2} a(sign \ Y_s) ds \cdot t + \varepsilon c_1 w_1(t)$$

$$\to t \cdot \int_{-\infty}^{\infty} a(y) \mu(dy) = t \cdot (a^- p^- + a^+ p^+), \tag{2.11}$$

where the limit is understood almost surely, in probability, and in L_1 [7], and the probabilities are defined as $p^+ = \mu(Y > 0)$ and $p^- = \mu(Y < 0)$ in a stationary regime. We can find them explicitly for they do not depend on the initial value $Y(0)$. So, let $Y(0) = \xi$ where ξ is a random variable with a distribution μ (of course, ξ does not depend on $w(s), \ s \ge 0$). Then

$$Y(s) = \xi + b^- \int_0^s I(Y(\theta) < 0) d\theta + b^+ \int_0^s I(Y(\theta) > 0) d\theta + c_2 w(s). \tag{2.12}$$

Due to the relation $EY(s) = E\xi, \ EI(Y(\theta) < 0) = p^-, \ EI(Y(\theta) > 0) = p^+$, we get

$$b^- p^- + b^+ p^+ = 0,$$

whence the relations (2.7) follow as $p^- + p^+ = 1$ and this implies also (2.6).

The invariant measure has on $(-\infty, 0)$ and on $(0, \infty)$ a density $p(y)$ which satisfies the following equations

$$\frac{1}{2} c_2^2 p'' - b^- \cdot p' = 0, \ y < 0, \ \int_{-\infty}^0 p(y) dy = p^-,$$

$$\frac{1}{2}c_2^2 p'' - b^+ \cdot p' = 0,\ y > 0,\ \int_0^\infty p(y)dy = p^+.$$

The relation (2.9) easily follows.

The assertions concerning the limit in $C([0,T];R^1)$ follow easily because the function a is bounded. Theorem 2.1 is proved.

2.2. General problem ($d = 2$). Now let us return to the original system (2.1). A perturbed system with small noise has the form (we observe that the result of Theorem 2.1 does not depend on c_1 and $c_2 \neq 0$; it turns out that the same is true in general case as well and we take $c_1 = c_2 = 1$ for simplicity in writing)

$$dX_t^\varepsilon = a(X_t^\varepsilon, Y_t^\varepsilon)dt + \varepsilon dw_1(t)$$

$$dY_t^\varepsilon = b(X_t^\varepsilon, Y_t^\varepsilon)dt + \varepsilon dw_2(t). \tag{2.13}$$

Remind that in contrast to (2.1), the system (2.13) has a pathwise unique strong solution, see [6].

Theorem 2.2. *Let $(x_0, y_0) \in S$ and*

$$(f^-, \nabla\varphi)_0 = a^-(x_0,y_0)\frac{\partial\varphi}{\partial x}(x_0,y_0) + b^-(x_0,y_0)\frac{\partial\varphi}{\partial y}(x_0,y_0) > 0,$$

$$(f^+, \nabla\varphi)_0 = a^+(x_0,y_0)\frac{\partial\varphi}{\partial x}(x_0,y_0) + b^+(x_0,y_0)\frac{\partial\varphi}{\partial y}(x_0,y_0) < 0, \tag{2.14}$$

where f is the vector with the components a, b and (remind) the vector $\nabla\varphi$ at any point $(x,y) \in S$ is directed to G^+. Let $(X_t^\varepsilon, Y_t^\varepsilon)$ be a solution of the system (2.13) starting from the point (x_0, y_0), $0 \le t \le \bar t$, where *$\bar t$ is some positive number. Then there exist the limits on $[0, \bar t\,]$ in probability, a.s., and in L_1*

$$\lim_{\varepsilon\to 0} X_t^\varepsilon := x(t),\ \lim_{\varepsilon\to 0} Y_t^\varepsilon := y(t), \tag{2.15}$$

where $x(0) = x_0,\ y(0) = y_0$ and $(x(t), y(t)) \in S,\ 0 \le t \le \bar t$.

The limit $(x(t), y(t))$ satisfies the system

$$\frac{dx(t)}{dt} = a^-(x(t),y(t))p^-(x(t),y(t)) + a^+(x(t),y(t))p^+(x(t),y(t)) := \breve a(x(t),y(t))$$

$$\frac{dy(t)}{dt} = b^-(x(t),y(t))p^-(x(t),y(t)) + b^+(x(t),y(t))p^+(x(t),y(t)) := \breve b(x(t),y(t)), \tag{2.16}$$

where

$$p^- = -\frac{(f^+,\nabla\varphi)}{(f^-,\nabla\varphi) - (f^+,\nabla\varphi)}\ ,\ p^+ = \frac{(f^-,\nabla\varphi)}{(f^-,\nabla\varphi) - (f^+,\nabla\varphi)}\ , \tag{2.17}$$

Proof. Our proof consists of two parts. The first one deals with deriving the law (2.16) for sliding mode provided its existence. Though the latter is evident from physical point of view, we prefer to give a complete proof in the second part as well. The parts are independent from each other.

Deriving the law (2.16). Consider a piece of the curve S. We assume that in some neighborhood of the point (x_0, y_0) it can be expressed by the equation $Y = \psi(X)$. Let us change the variables X, Y :

$$\tilde X = X,\ \tilde Y = Y - \psi(X). \tag{2.18}$$

We have

$$d\tilde{X}_t^\varepsilon = a(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon + \psi(\tilde{X}_t^\varepsilon))dt + \varepsilon dw_1(t) := \tilde{a}(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon)dt + \varepsilon dw_1(t)$$

$$d\tilde{Y}_t^\varepsilon = b(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon + \psi(\tilde{X}_t^\varepsilon))dt$$

$$-\psi'(\tilde{X}_t^\varepsilon)\cdot(a(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon + \psi(\tilde{X}_t^\varepsilon))dt + \varepsilon dw_1(t)) - \frac{1}{2}\varepsilon^2\psi''(\tilde{X}_t^\varepsilon)dt + \varepsilon dw_2(t)$$

$$:= \tilde{b}(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon)dt - \frac{1}{2}\varepsilon^2\psi''(\tilde{X}_t^\varepsilon)dt - \varepsilon\psi'(\tilde{X}_t^\varepsilon)dw_1(t) + \varepsilon dw_2(t). \tag{2.19}$$

Consider a δ-neighborhood $\tilde{U}_\delta$ of the point $(x_0, 0)$: $\tilde{U}_\delta = \{(\tilde{X},\ \tilde{Y}) :\ |\tilde{X} - x_0| < \delta,\ |\tilde{Y}| < \delta\}$ with small δ. Let $(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon)$ be the solution of the system (2.19) starting from the point $(x_0, 0)$. If δ and ε are sufficiently small then the exit probability of this solution from $\tilde{U}_\delta$ during the time $\Delta = \delta^{1/\alpha}$, $0 < \alpha < 1$, is very small. Strictly speaking the following is true for the stopped on $\partial\tilde{U}_\delta$ process. But in order to avoid unessential complications we consider the process $(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon)$ to belong to $\tilde{U}_\delta$ during the time Δ (and even during a little bigger time $\Delta + O(\Delta^{1+\alpha})$).

Let us consider the time change which makes the diffusion coefficient of $\tilde{Y}_t^\varepsilon$ constant. Namely, we introduce the new (random) time $\bar{t} = \bar{t}(t)$:

$$\bar{t}(t) := \int_0^t (1 + \psi'(\tilde{X}_\theta^\varepsilon)^2)\cdot(1 + \psi'(x_0)^2)^{-1}d\theta. \tag{2.20}$$

Clearly, if $0 \le t \le \Delta$, then $0 \le \bar{t} \le \Delta + O(\Delta^{1+\alpha})$ and, vice versa, if $0 \le \bar{t} \le \Delta$, then $0 \le t \le \Delta + O(\Delta^{1+\alpha})$. In particular

$$t(\Delta) = \Delta + O(\Delta^{1+\alpha}). \tag{2.21}$$

Denote $(\bar{X}_{\bar{t}}^\varepsilon, \bar{Y}_{\bar{t}}^\varepsilon) = (\tilde{X}_{t(\bar{t})}^\varepsilon, \tilde{Y}_{t(\bar{t})}^\varepsilon)$ on $[0, \Delta]$. We get

$$d\bar{X}_{\bar{t}}^\varepsilon = \bar{a}(\bar{X}_{\bar{t}}^\varepsilon, \bar{Y}_{\bar{t}}^\varepsilon)d\bar{t} + \varepsilon\frac{(1 + \psi'(x_0)^2)^{1/2}}{(1 + \psi'(\bar{X}_{\bar{t}}^\varepsilon)^2)^{1/2}}d\bar{w}_1(\bar{t})$$

$$d\bar{Y}_{\bar{t}}^\varepsilon = \bar{b}(\bar{X}_{\bar{t}}^\varepsilon, \bar{Y}_{\bar{t}}^\varepsilon)d\bar{t} - \frac{\varepsilon^2}{2}\psi''(\bar{X}_{\bar{t}}^\varepsilon)\frac{1 + \psi'(x_0)^2}{1 + \psi'(\bar{X}_{\bar{t}}^\varepsilon)^2}d\bar{t} + \varepsilon\sqrt{1 + \psi'(x_0)^2}d\bar{w}_2(\bar{t}), \tag{2.22}$$

where $\bar{w}_1,\ \bar{w}_2$ are new standard Wiener processes (they are dependent, however, it does not matter), and

$$\bar{a}(x, y) = \tilde{a}(x, y)\cdot\frac{1 + \psi'(x_0)^2}{1 + \psi'(x)^2},\ \bar{b}(x, y) = \tilde{b}(x, y)\cdot\frac{1 + \psi'(x_0)^2}{1 + \psi'(x)^2}\ .$$

Clearly

$$\bar{a}^\pm(x_0, 0) = a^\pm(x_0, y_0) := \bar{a}_0^\pm\ ,$$

$$\bar{b}^\pm(x_0, 0) = -\psi'(x_0)a^\pm(x_0, y_0) + b^\pm(x_0, y_0) := \bar{b}_0^\pm\ .$$

Due to (2.14) and because the vector $(-\psi'(x_0), 1)$ is collinear to the vector $\nabla\varphi(x_0, y_0)$ we have

$$\bar{b}_0^- > 0,\ \bar{b}_0^+ < 0. \tag{2.23}$$

It is not difficult to show that

$$\bar{a}(x, y) = a_0(sign\ y) + \alpha(x, y),\ \bar{b}(x, y) = b_0(sign\ y) + \beta(x, y),$$

where

$$a_0(sign\ y) = \begin{cases} \bar{a}_0^-, \ y<0, \\ 0, \ y=0, \\ \bar{a}_0^+, \ y>0, \end{cases}, \ b_0(sign\ y) = \begin{cases} \bar{b}_0^-, \ y<0, \\ 0, \ y=0, \\ \bar{b}_0^+, \ y>0, \end{cases},$$

and there exists a constant $K>0$ such that

$$|\alpha(x,y)| \le K(|x-x_0|+|y|), \ |\beta(x,y)| \le K(|x-x_0|+|y|). \tag{2.24}$$

Introduce the process

$$Z_s = \varepsilon^{-2}\bar{Y}^\varepsilon_{\varepsilon^2 s}\ , \ 0 \le s \le \frac{\Delta}{\varepsilon^2}\ . \tag{2.25}$$

The process Z_s depends on ε. However the dependence is not too essential and we do not mark it. We have

$$dZ_s = b_0(sign\ Z_s)ds + \beta(\bar{X}^\varepsilon_{\varepsilon^2 s}, \bar{Y}^\varepsilon_{\varepsilon^2 s})ds - \frac{\varepsilon^2}{2}\psi''(\bar{X}^\varepsilon_{\varepsilon^2 s})\frac{1+\psi'(x_0)^2}{1+\psi'(\bar{X}^\varepsilon_{\varepsilon^2 s})^2}ds$$

$$+\sqrt{1+\psi'(x_0)^2}dw(s), \ Z_0 = 0, \ 0 \le s \le \frac{\Delta}{\varepsilon^2}\ , \tag{2.26}$$

where $w(s) = \varepsilon^{-1}\bar{w}_2(\varepsilon^2 s)$.

Since $(\bar{X}^\varepsilon_{\varepsilon^2 s}, \bar{Y}^\varepsilon_{\varepsilon^2 s}) \in \tilde{U}_\delta$, we get due to (2.24) (we consider $\varepsilon < \Delta^{\alpha/2}$ and we note that various constants are given by the same letter K)

$$|\beta(\bar{X}^\varepsilon_{\varepsilon^2 s}, \bar{Y}^\varepsilon_{\varepsilon^2 s}) - \frac{\varepsilon^2}{2}\psi''(\bar{X}^\varepsilon_{\varepsilon^2 s})\frac{1+\psi'(x_0)^2}{1+\psi'(\bar{X}^\varepsilon_{\varepsilon^2 s})^2}| \le K\Delta^\alpha.$$

Introduce another two processes $\check{Z}_s$ and $\hat{Z}_s$ which solve the equations

$$d\check{Z}_s = b_0(sign\ \check{Z}_s)ds - K\Delta^\alpha ds + \sqrt{1+\psi'(x_0)^2}dw(s), \ \check{Z}_0 = 0,$$

$$d\hat{Z}_s = b_0(sign\ \hat{Z}_s)ds + K\Delta^\alpha ds + \sqrt{1+\psi'(x_0)^2}dw(s), \ \hat{Z}_0 = 0.$$

Due to the comparison theorem [2] (we note that the comparison theorem in [2] is proved under some other conditions but it can be carried over to the considered case)

$$\check{Z}_s \le Z_s \le \hat{Z}_s.$$

The invariant measures $\check{\mu}$ and $\hat{\mu}$ of the processes have the density of the form (2.9) with

$$\check{p}^- = \frac{\bar{b}_0^+ - K\Delta^\alpha}{\bar{b}_0^+ - \bar{b}_0^-}, \ \check{p}^+ = \frac{\bar{b}_0^- - K\Delta^\alpha}{\bar{b}_0^- - \bar{b}_0^+}\ ,$$

$$\hat{p}^- = \frac{\bar{b}_0^+ + K\Delta^\alpha}{\bar{b}_0^+ - \bar{b}_0^-}, \ \hat{p}^+ = \frac{\bar{b}_0^- + K\Delta^\alpha}{\bar{b}_0^- - \bar{b}_0^+}\ .$$

Clearly

$$\check{p}^\pm = \pm\frac{(f^\mp, \nabla\varphi)}{(f^-, \nabla\varphi) - (f^+, \nabla\varphi)} + O(\Delta^\alpha) = p^\pm + O(\Delta^\alpha).$$

Analogously

$$\hat{p}^\pm = p^\pm + O(\Delta^\alpha).$$

If $\bar{a}_0^+ \ge \bar{a}_0^-$, then

$$a_0(sign\ \check{Z}_s) \le a_0(sign\ \bar{Y}^\varepsilon_{\varepsilon^2 s}) = a_0(sign\ Z_s) \le a_0(sign\ \hat{Z}_s).$$

If $\bar{a}_0^- \geq \bar{a}_0^+$, then the previous inequality is replaced by the contrary one. It is not difficult to prove (as in Theorem 2.1) that in both cases there exist the limits

$$\liminf_{\varepsilon\to 0} \bar{X}^\varepsilon_\Delta = \Delta \cdot (\bar{a}_0^- p^- + \bar{a}_0^+ p^+) + O(\Delta^{1+\alpha}),$$

$$\limsup_{\varepsilon\to 0} \bar{X}^\varepsilon_\Delta = \Delta \cdot (\bar{a}_0^- p^- + \bar{a}_0^+ p^+) + O(\Delta^{1+\alpha}). \tag{2.27}$$

As $\bar{X}^\varepsilon_\Delta = \tilde{X}^\varepsilon_{t(\Delta)}$, we have (due to (2.21)) the same relations for $\tilde{X}^\varepsilon_\Delta$ as well. Using (2.25), we get

$$\lim_{\varepsilon\to 0} \tilde{Y}^\varepsilon_\Delta = 0. \tag{2.28}$$

Returning to the original variables, we obtain from (2.27) (for $\tilde{X}^\varepsilon_\Delta$) and (2.28) that

$$\lim_{\Delta\to 0} \frac{1}{\Delta} \liminf_{\varepsilon\to 0} (X^\varepsilon_\Delta - x_0) = \lim_{\Delta\to 0} \frac{1}{\Delta} \limsup_{\varepsilon\to 0} (X^\varepsilon_\Delta - x_0) = \breve{a}(x_0, y_0),$$

$$\lim_{\Delta\to 0} \frac{1}{\Delta} \liminf_{\varepsilon\to 0} (Y^\varepsilon_\Delta - y_0) = \lim_{\Delta\to 0} \frac{1}{\Delta} \limsup_{\varepsilon\to 0} (Y^\varepsilon_\Delta - y_0)$$

$$= \lim_{\Delta\to 0} \frac{1}{\Delta} \liminf_{\varepsilon\to 0} (\psi(X^\varepsilon_\Delta) - y_0) = \lim_{\Delta\to 0} \frac{1}{\Delta} \limsup_{\varepsilon\to 0} (\psi(X^\varepsilon_\Delta) - y_0) = \psi'(x_0)\breve{a}(x_0, y_0).$$

The equality

$$\psi'(x_0)\breve{a}(x_0, y_0) = \breve{b}(x_0, y_0)$$

follows due to the relation (see formulae (2.17))

$$(f^-, \nabla\varphi)p^- + (f^+, \nabla\varphi)p^+ = 0.$$

So we prove that the limit process defines the field of vectors $(\breve{a}, \breve{b})$ on the curve S which acts tangentially to S and, consequently, the limit process satisfies the system (2.16). In the capacity of $\bar{t}$ one can take a time on which the solution of the system (2.16) with initial data $x(0) = x_0,\ y(0) = y_0$ exists.

The complete proof. It is convenient to divide this part of the proof into steps.

- Let us take a partition of the interval $[0, T]$ by the points $t_k = k\Delta,\ k = 0, 1, \ldots, N$. Assume that Δ is small. Notice that both $\tilde{X}^\varepsilon_s$ and $\tilde{Y}^\varepsilon_s$ are close to its values in t_k in probability as s is close to t_k, i.e. for any $c > 0$ one has,

$$\sup_{\varepsilon\leq 1} P(\sup_{k\leq N} \sup_{|s-t_k|\leq\Delta} \|(\tilde{X}^\varepsilon_s, \tilde{Y}^\varepsilon_s) - (\tilde{X}^\varepsilon_{t_k}, \tilde{Y}^\varepsilon_{t_k})\| > c) = o_\Delta(1).$$

- Let us consider the time change which makes the diffusion coefficient of $\tilde{Y}^\varepsilon$ equal to *const* times ε and the diffusion coefficient of $\tilde{X}^\varepsilon$ close to another *const* times ε on each partition interval $[t_k, t_{k+1}]$, namely, let

$$h_t := \int_0^t (1 + \psi'(\tilde{X}^\varepsilon_s)^2)(1 + \psi'(\tilde{X}^\varepsilon_{[s/\Delta]\Delta})^2)^{-1}\, ds,$$

and new time $t' = t'(t) = h_t^{-1}$. Notice that dh_t/dt is close to 1 in probability if Δ is small enough: for any $t > 0$ and any $c > 0$

$$\sup_{\varepsilon\leq 1} P(\sup_{s\leq t} |\dot{h}_s - 1| > c) = o_\Delta(1).$$

- Denote $(\tilde X'_t, \tilde Y'_t) = (\tilde X^\varepsilon_{t'(t)}, \tilde Y^\varepsilon_{t'(t)})$ (we omit the index ε for the notation simplicity). Then (see [2])

$$d\tilde X'_t = \hat a_t(\tilde X'_t, \tilde Y'_t)\,dt + \varepsilon(1+\psi'(\tilde X'_{[t/\Delta]\Delta})^2)^{1/2}(1+\psi'(\tilde X'_t))^{-1/2}\,d\bar w_1(t),$$

$$d\tilde Y'_t = \hat b_t(\tilde X'_t, \tilde Y'_t)\,dt + \varepsilon(1+\psi'(\tilde X'_{[t/\Delta]\Delta}))^{1/2}\,d\bar w(t),$$

i.e., on each partition interval $[t_k, t_{k+1}[$,

$$d\tilde X'_t = \hat a_t(\tilde X'_t, \tilde Y'_t)\,dt + \varepsilon(1+\psi'(\tilde X'_{t_k})^2)^{1/2}(1+\psi'(\tilde X'_t))^{-1/2}\,d\bar w_1(t),$$

$$d\tilde Y'_t = \hat b_t(\tilde X'_t, \tilde Y'_t)\,dt + \varepsilon(1+\psi'(\tilde X'_{t_k}))^{1/2}\,d\bar w(t),$$

where $\bar w_1$ and $\bar w$ are new standard Wiener processes (they are dependent but it does not matter) and the coefficients $\hat a$ and $\hat b$ have the form

$$\hat a_t(\tilde x', \tilde y') = \tilde a(\tilde x', \tilde y')(1+\psi'(\tilde X'_{[t/\Delta]\Delta})^2)(1+\psi'(\tilde x')^2)^{-1},$$

$$\hat b_t(\tilde x', \tilde y') = \left(\tilde b(\tilde x', \tilde y') - \varepsilon^2\psi''(\tilde x')/2\right)(1+\psi'(\tilde X'_{[t/\Delta]\Delta})^2)(1+\psi'(\tilde x')^2)^{-1}.$$

- It is easy to see that $\sup_{s\le t}|\tilde Y_s| \to 0$ in probability as $\varepsilon\to 0$ for any $t\le T$.
- Further, let us consider a new process

$$\tilde X''_t = X_0 + \int_0^t \hat a_s(\tilde X'_{[s/\Delta]\Delta}, \tilde Y'_s)\,ds.$$

  Evidently, $\tilde X'_t = \tilde X''_t + o_{\Delta,\varepsilon}(1)$.
- Let $\hat a^+_t(x) = \lim\limits_{y\downarrow 0}\hat a_t(x,y)$, $\hat a^-_t(x) = \lim\limits_{y\uparrow 0}\hat a_t(x,y)$. Since $\sup_{s\le t}|\tilde Y_s| \to 0$ as $\varepsilon\to 0$ then

$$\begin{aligned}\tilde X''_t &= X_0 + \sum_{k\le N}\Bigg\{\hat a^+_{k\Delta}(\tilde X'_{k\Delta})\Delta^{-1}\int\limits_{k\Delta}^{(k+1)\Delta} 1(\tilde Y'_s>0)\,ds \\ &\quad + \hat a^-_{k\Delta}(X'_{k\Delta})\Delta^{-1}\int\limits_{k\Delta}^{(k+1)\Delta} 1(\tilde Y'_s<0)\,ds\Bigg\} + o_\varepsilon(1),\end{aligned}$$

where $(P)\lim_{\varepsilon\to 0} o_\varepsilon(1) = 0$.

  Denote $\nu'_k = \Delta^{-1}\int\limits_{k\Delta}^{(k+1)\Delta} 1(Y'_s>0)\,ds$. Then one can rewrite the last assertion as

$$\tilde X''_t = X_0 + \sum_{k\le N}\{\hat a^+_{k\Delta}(\tilde X'_{k\Delta})\nu'_k + \hat a^-_{k\Delta}(\tilde X'_{k\Delta})(1-\nu'_k)\} + o_\varepsilon(1).$$

- Let $U(x,0) = \{(x',y') : \ |x-x'|\le r,\ |y'|\le r\}$, r is small and $\Delta << r$. We introduce new random coefficients which also depend on time,

$$\begin{aligned}\hat b_{sup}(t,x,y) &= \sup\{\hat b(x',y') :\ (x',y')\in U(\tilde X'_{[t/\Delta]\Delta},0), sign\,y' = sign\,y\} + \delta, \\ \hat b_{inf}(x,y) &= \inf\{\hat b(x',y') :\ (x',y')\in U(\tilde X'_{[t/\Delta]\Delta},0), sign\,y' = sign\,y\} + \delta,\end{aligned}$$

where $\delta>0$ is one more small parameter.

Notice that $P\left(\hat b_{sup}(t,x,y) \approx \hat b(\tilde X'_{[t/\Delta]\Delta}, y) \approx \hat b_{inf}(t,x,y),\ 0 \le t \le T\right) \approx 1$ if r and δ are small enough. Consider new processes which solve the equations

$$
\begin{aligned}
Y''^{sup}_t &= Y_0 + \int_0^t \hat b_{sup}(\tilde X'_{[s/\Delta]\Delta}, Y''^{sup}_s)\, ds + \varepsilon \int_0^t (1 + \psi'(\tilde X'_{[s/\Delta]\Delta}))^{1/2}\, d\bar w_s, \\
Y''^{inf}_t &= Y_0 + \int_0^t \hat b_{inf}(\tilde X'_{[s/\Delta]\Delta}, Y''^{inf}_s)\, ds + \int_0^t \varepsilon (1 + \psi'(\tilde X'_{[s/\Delta]\Delta}))^{1/2} d\bar w_s,
\end{aligned}
$$

Denote $U_k = U(X_{k\Delta}, 0)$, $T_k := \inf(t \ge k\Delta :\ (\tilde X, \tilde Y) \notin U_k)$. Due to comparison theorems,

$$
P(Y''^{inf}_{\min(t,T_k)} \le Y'_{\min(t,T_k)} \le Y''^{sup}_{\min(t,T_k)}) = 1.
$$

Indeed, it follows from arguments of ODE theory since the diffusion coefficients of both processes are equal and piecewise constant.

- We can write

$$
P(\nu''^{inf}_k \le \nu'_k \le \nu''^{sup}_k,\ k \le N) \approx 1
$$

where

$$
\nu''^{sup} = \Delta^{-1} \int_{k\Delta}^{(k+1)\Delta} 1(Y''^{sup}_s)\, ds, \quad \nu''^{inf} = \Delta^{-1} \int_{k\Delta}^{(k+1)\Delta} 1(Y''^{inf}_s)\, ds.
$$

- Now, we make the change of space and time exactly as in the proof of Theorem 2.1. Due to the convergence to the invariant measure, we get on each partition interval $k\Delta, (k+1)\Delta$ given $\tilde X'_{k\Delta}$,

$$
\nu''^{inf}_k - \frac{\hat b^-_{inf}(\tilde X'_{k\Delta})}{\hat b^-_{inf}(\tilde X'_{k\Delta}) - \hat b^+_{inf}(\tilde X'_{k\Delta})} \to 0
$$

and

$$
\nu''^{sup}_k - \frac{\hat b^-_{sup}(\tilde X'_{k\Delta})}{\hat b^-_{sup}(\tilde X'_{k\Delta}) - \hat b^+_{sup}(\tilde X'_{k\Delta})} \to 0.
$$

Since

$$
\frac{\hat b^-_{inf}(\tilde X'_{k\Delta})}{\hat b^-_{inf}(\tilde X'_{k\Delta}) - \hat b^+_{inf}(\tilde X'_{k\Delta})} \approx \frac{\hat b^-_{sup}(\tilde X'_{k\Delta})}{\hat b^-_{sup}(\tilde X'_{k\Delta}) - \hat b^+_{sup}(\tilde X'_{k\Delta})}
$$

(see Theorem 2.1), we get

$$
\nu''^{inf}_k \approx \nu''^{sup}_k, \quad \text{for all } k.
$$

- Hence,

$$
\begin{aligned}
\tilde X'_t - x &= \sum_k \int_{k\Delta}^{(k+1)\Delta} \hat a(\tilde X'_{t_k}, y) \mu^{sup}(\tilde X'_{t_k}, dy) + o(1) \\
&= \sum_k \int_{k\Delta}^{(k+1)\Delta} \hat a(\tilde X'_{t_k}, y) \mu^{inf}(\tilde X'_{t_k}, dy) + o(1),
\end{aligned}
$$

where μ^{sup} and μ^{inf} are invariant measures of the processes Y''^{sup} and Y''^{inf} correspondingly and $o(1) \to 0$ in probability as Δ, δ, $\varepsilon \to 0$.

Finally, since the measures $\mu^{sup}(x, dy)$ and $\mu^{inf}(x, dy)$ are close to $\mu(x, dy)$ and the last measure is continuous in x (in the weak sense), we get

$$\tilde{X}'_t - x = \int_0^t \hat{a}(\tilde{X}'_s, y)\mu(\tilde{X}'_s, dy) + o(1).$$

Here $o(1) \to 0$ in probability as $\varepsilon \to 0$.

- Remind that the time change $t'(t)$ is such that $dt'(t)/dt$ is uniformly close to 1 in probability if Δ is small enough. As $\Delta \to 0$, we obtain

$$\tilde{X}_t - x = \int_0^t \tilde{a}(\tilde{X}_s, y)\mu(\tilde{X}_s, dy) + o(1).$$

This is equivalent to the desired assertion. Theorem 2.2 is proved.

Remark 2.1. The established law (2.16) of motion on S coincides with the well known sliding mode [1].

Remark 2.2. Clearly, the vector $(\breve{a}(x, y), \breve{b}(x, y))$, $(x, y) \in S$, is tangent to S at the point (x, y). After the change of variables it coincides with the vector $(\bar{a}, 0)$ from Theorem 2.1 (of course, the probabilities $p^-(x, y)$, $p^+(x, y)$ from (2.17) are equal to p^-, p^+ from (2.7)). Thus, Theorem 2.2 justifies the following principle: to obtain the infinitesimal characteristics of motion on a sliding surface one should reduce the problem to the corresponding model one. The system (2.5) gives an example of such a model problem.

Remark 2.3. The system (2.16) can be rewritten in the form

$$\frac{dx(t)}{dt} = f^-(x(t))p^-(x(t)) + f^+(x(t))p^+(x(t)) := \bar{f}(x(t)). \tag{2.29}$$

We remind that f is the vector with the components a, b.

2.3. Model and general d-dimensional problems. The obtained results can be generalized to the case of a d-dimensional system (1.1) if the surface of discontinuity S is $(d-1)$-dimensional. Slightly abusing initial notation, we can represent the system (1.1) in the form

$$\frac{dx}{dt} = a(x, y), \ \frac{dy}{dt} = b(x, y), \tag{2.30}$$

where $x = (x^1, ..., x^{d-1})$, $y = x^d$.

Let $(x_0, y_0) = (x_0^1, ..., x_0^{d-1}, y_0) \in S$ be a fixed point. Assume that the surface S can be expressed by the equation $y = \psi(x^1, ..., x^{d-1})$ in a neighborhood of (x_0, y_0). Introduce G^+ by the rule: the points (x, y) with $y > \psi(x)$ from a neighborhood of (x_0, y_0) belong to G^+. Suppose that

$$b^-(x_0, y_0) - (\nabla\psi(x_0), a^-(x_0, y_0)) > 0, \ b^+(x_0, y_0) - (\nabla\psi(x_0), a^+(x_0, y_0)) < 0. \tag{2.31}$$

Consider new coordinates

$$X = x - x_0, \ Y = y - \psi(x).$$

Then

$$\frac{dX}{dt} = a(X + x_0, Y + \psi(X + x_0)) := \tilde{a}(X, Y)$$

$$\frac{dY}{dt} = b(X + x_0, Y + \psi(X + x_0))$$

$$-(\nabla\psi(X + x_0), a(X + x_0, Y + \psi(X + x_0))) := \tilde{b}(X, Y). \tag{2.32}$$

For this system the surface of discontinuity is expressed by the equation $Y = 0$ in a neighborhood of the origin, the domain G^+ is the upper half-space $Y > 0$ and the domain G^- is the lower half-space $Y < 0$, and due to (2.31) $\tilde{b}^-(0,0) > 0$, $\tilde{b}^+(0,0) < 0$. The same arguments, which were used for deriving the system (2.3), lead to the system

$$\dot{X} = a(Y),\ \dot{Y} = b(Y), \tag{2.33}$$

where (as before $a(Y) = a(sign\ Y)$, $b(Y) = b(sign\ Y)$)

$$a(Y) = \begin{cases} a^-,\ Y < 0, \\ 0,\ Y = 0, \\ a^+,\ Y > 0, \end{cases},\ b(Y) = \begin{cases} b^-,\ Y < 0, \\ 0,\ Y = 0, \\ b^+,\ Y > 0, \end{cases} \tag{2.34}$$

with the vectors $a^- = \tilde{a}^-(0,0)$, $a^+ = \tilde{a}^+(0,0)$, and with the scalars $b^- = \tilde{b}^-(0,0) > 0$, $b^+ = \tilde{b}^+(0,0) < 0$.

Along with the system (2.33) consider the d-dimensional system with small noise

$$dX_t^\varepsilon = a(Y_t^\varepsilon)dt + \varepsilon dw_1(t)$$

$$dY_t^\varepsilon = b(Y_t^\varepsilon)dt + \varepsilon dw_2(t), \tag{2.35}$$

where w_1 and w_2 are $(d-1)$-dimensional and one-dimensional independent standard Wiener processes correspondingly.

For the system (2.35) one can prove the same result as Theorem 2.1 asserts. Namely, let $(X_t^\varepsilon, Y_t^\varepsilon)$ be the solution of the system (2.35) starting from the origin. Then its limit as $\varepsilon \to 0$ equals

$$\lim_{\varepsilon\to 0} X_t^\varepsilon = \bar{a}\cdot t,\ \lim_{\varepsilon\to 0} Y_t^\varepsilon = 0,$$

where the $(d-1)$-dimensional vector $\bar{a}$ is equal to $a^-p^- + a^+p^+$ with p^-, p^+ from (2.7).

A theorem analogous to Theorem 2.2 can be proved as well.

Let S be a surface of discontinuity for the system (1.1) expressed by the equation $\varphi(x) = 0$. Remember, the domain G^+ is chosen so that the vector $\nabla\varphi(x)$ at the point $x \in S$ is directed to it.

Along with the system (2.30) let us consider a perturbed system

$$dX_t^\varepsilon = f(X_t^\varepsilon)dt + \varepsilon dw(t), \tag{2.36}$$

where $w(t)$ is a d-dimensional standard Wiener process.

Theorem 2.3. *Let*

$$(f^-, \nabla\varphi) > 0,\ (f^+, \nabla\varphi) < 0. \tag{2.37}$$

and X_t^ε be a solution of the system (2.36) starting from a point $x \in S$, $0 \le t \le \bar{t}$, where $\bar{t}$ is a positive number. Then there exist the limits on $[0, \bar{t}\,]$ in probability, a.s., and in L_1

$$\lim_{\varepsilon\to 0} X_t^\varepsilon := x(t),$$

where $x(0) = x$ and $x(t) \in S,\ 0 \le t \le \bar{t}$.

The limit $x(t)$ satisfies the system

$$\frac{dx(t)}{dt} = f^-(x(t))p^-(x(t)) + f^+(x(t))p^+(x(t)) := \bar{f}(x(t)) \tag{2.38}$$

with

$$p^-(x) = -\frac{(f^+, \nabla\varphi)}{(f^-, \nabla\varphi) - (f^+, \nabla\varphi)}\ ,\ p^+(x) = \frac{(f^-, \nabla\varphi)}{(f^-, \nabla\varphi) - (f^+, \nabla\varphi)}. \tag{2.39}$$

3. Sliding mode under weak and strong singularity

Consider the following two-dimensional model system

$$\dot{X} = a(Y),\ \dot{Y} = |Y|^\gamma b(Y),\ |\gamma| < 1, \tag{3.1}$$

with a and b from (2.4).

Under $\gamma = 0$ we get the system (2.3). All the trajectories of the system (2.3) approach the line of discontinuity $Y = 0$ for a finite time and arrive at this line with a nonzero (because $b^- \neq 0,\ b^+ \neq 0$) and bounded Y-component of velocity. If $\gamma \neq 0$, as before they approach the line of discontinuity for a finite time but with zero $(0 < \gamma < 1)$ or with infinite $(-1 < \gamma < 0)$ final Y-component of velocity. Therefore the obtained rule of sliding mode cannot be used.

Consider the system with small noise

$$dX_t^\varepsilon = a(Y_t^\varepsilon)dt + \varepsilon dw_1(t)$$

$$dY_t^\varepsilon = |Y_t^\varepsilon|^\gamma b(Y_t^\varepsilon)dt + \varepsilon dw_2(t). \tag{3.2}$$

Introduce the process

$$Y_s = \varepsilon^{-\alpha} Y_{\varepsilon^\beta s}^\varepsilon\ ,\ \alpha = \frac{2}{1+\gamma}\ ,\ \beta = \frac{2(1-\gamma)}{1+\gamma}. \tag{3.3}$$

It is not difficult to show that the law of the process Y_s does not depend on ε and Y_s satisfies the equation

$$dY = |Y|^\gamma b(Y)ds + dw(s) \tag{3.4}$$

with the standard Wiener process $w(s) = \varepsilon^{-\beta/2} w_2(\varepsilon^\beta s)$.

The Markov process defined by the stochastic differential equation (3.4) is ergodic (see [8]). Its invariant measure μ has on $(-\infty, 0)$ and on $(0, \infty)$ a density $p(y)$ which satisfies the following equations

$$\frac{1}{2}p'' - b^- \cdot \frac{\partial(|y|^\gamma p)}{\partial y} = 0,\ y < 0,$$

$$\frac{1}{2}p'' - b^+ \cdot \frac{\partial(y^\gamma p)}{\partial y} = 0,\ y > 0. \tag{3.5}$$

We calculate

$$p(y) = \begin{cases} C_1 \exp(-\dfrac{b^-}{\gamma+1}|y|^{\gamma+1})\ ,\ y < 0, \\ \\ C_2 \exp(\dfrac{b^+}{\gamma+1}y^{\gamma+1})\ ,\ y > 0. \end{cases} \tag{3.6}$$

One can show the equality $C_1 = C_2$ as follows. From (3.4) we have

$$Y(s) = \xi + \int_0^s |Y(\theta)|^\gamma b(Y(\theta))d\theta + w(s),$$

where ξ has the distribution (3.6).

Since $EY(s) = E\xi$, we get

$$E|Y(s)|^\gamma b(Y(s)) = \int_{-\infty}^0 b^- \cdot (-y)^\gamma p(y)dy + \int_0^\infty b^+ \cdot y^\gamma p(y)dy = 0,$$

whence the equality $C_1 = C_2 := C$ can be obtained easily.

Further, the first equation of (3.2) gives (cf. (2.11))

$$X^\varepsilon_{\Delta t} = \int_0^{\Delta t} a(Y^\varepsilon_t)dt + \varepsilon w_1(\Delta t)$$

$$= \int_0^{\Delta t} a(\varepsilon^\alpha Y_{t/\varepsilon^\beta})dt + \varepsilon w_1(\Delta t) = \int_0^{\Delta t} a(sign\ Y_{t/\varepsilon^\beta})dt + \varepsilon w_1(\Delta t)$$

$$= \frac{\varepsilon^\beta}{\Delta t} \int_0^{\Delta t/\varepsilon^\beta} a(sign\ Y_s)ds \cdot \Delta t + \varepsilon w_1(\Delta t)$$

$$\to \Delta t \cdot \int_{-\infty}^\infty a(y)\mu(dy) = \Delta t \cdot (a^- p^- + a^+ p^+).$$

Let us find the values p^- and p^+. We have

$$p^- = \int_{-\infty}^0 p(y)dy = C\int_0^\infty \exp(-\frac{b^-}{\gamma+1}y^{\gamma+1})dy = \frac{C}{(b^-)^{1/\gamma}}\int_0^\infty \exp(-\frac{y^{\gamma+1}}{\gamma+1})dy,$$

$$p^+ = \int_0^\infty p(y)dy = C\int_0^\infty \exp(\frac{b^+}{\gamma+1}y^{\gamma+1})dy = \frac{C}{(-b^+)^{1/\gamma}}\int_0^\infty \exp(-\frac{y^{\gamma+1}}{\gamma+1})dy.$$

Since $p^- + p^+ = 1$, we get

$$p^- = \frac{(-b^+)^{1/\gamma}}{(b^-)^{1/\gamma} + (-b^+)^{1/\gamma}}\ ,\ p^+ = \frac{(b^-)^{1/\gamma}}{(b^-)^{1/\gamma} + (-b^+)^{1/\gamma}}\ .$$

Thus, the sliding mode on the sliding line $Y = 0$ for the system (3.1) is the uniform motion with the speed $a^- p^- + a^+ p^+$.

4. Stochastic sliding mode on $(d-1)$-dimensional surfaces of discontinuity

4.1. Model problem. Consider the following d-dimensional stochastic system

$$dX = a(Y)dt + \sigma(Y)dW(t)$$

$$dY = b(Y)dt, \tag{4.1}$$

where X and a are $(d-1)$-dimensional vectors, Y and b are scalars, W is a k-dimensional standard Wiener process, σ is a $(d-1)\times k$ matrix. We suppose that the coefficients $a,\ \sigma,\ b$ depend on *sign* Y only:

$$a(Y) = \begin{cases} a^-,\ Y<0, \\ 0,\ Y=0, \\ a^+,\ Y>0, \end{cases},\ \sigma(Y) = \begin{cases} \sigma^-,\ Y<0, \\ 0,\ Y=0, \\ \sigma^+,\ Y>0, \end{cases},\ b(Y) = \begin{cases} b^-,\ Y<0, \\ 0,\ Y=0, \\ b^+,\ Y>0, \end{cases} \tag{4.2}$$

and that

$$b^- > 0,\ b^+ < 0.$$

Thus, we have two media ($Y < 0$ and $Y > 0$) with two different laws of stochastic motion. In addition, any solution approaches the plane of discontinuity $Y = 0$ for a finite time and cannot be able to leave this plane. Our goal is to obtain a "natural" law of stochastic sliding mode on sliding plane induced by the stochastic motion in these two media.

To this end introduce the following d-dimensional system with additional small noise

$$dX_t^\varepsilon = a(Y_t^\varepsilon)dt + \sigma(Y_t^\varepsilon)dW(t) + \varepsilon dw_1(t)$$

$$dY_t^\varepsilon = b(Y_t^\varepsilon)dt + \varepsilon c dw_2(t), \tag{4.3}$$

where w_1 and w_2 are $(d-1)$-dimensional and one-dimensional independent standard Wiener processes correspondingly. The processes w_1 and w_2 are independent of W as well.

We define the sliding diffusion as a limit process for X_t^ε under ε tending to zero provided the process $(X_t^\varepsilon,\ Y_t^\varepsilon)$ starts from the origin.

Theorem 4.1. *The processes $(X_t^\varepsilon,\ 0 \le t \le T)$ tend weakly in distribution to a Gaussian process*

$$\bar{X}_t = \bar{a}\, t + \bar{\sigma}\, \bar{W}_t,$$

where p^- and p^+ are from (2.7),

$$\bar{a} = a^- p^- + a^+ p^+, \quad \bar{\sigma} = (\sigma^-(\sigma^-)^T p^- + \sigma^+(\sigma^+)^T p^+)^{1/2}$$

(the square root of the non-negative definite constant symmetric matrix is always well-defined), and $\bar{W}$ is a $(d-1)$-dimensional standard Wiener process.

Proof. As earlier introduce the process $Y_s = \varepsilon^{-2} Y^\varepsilon_{\varepsilon^2 s}$. We have for any $t > 0$ (cf. (2.11))

$$X_t^\varepsilon = \int_0^t a(sign\ Y_{s/\varepsilon^2})ds + \int_0^t \sigma(sign\ Y_{s/\varepsilon^2})dW(s) + \varepsilon w_1(t)$$

$$= \varepsilon^2 \int_0^{t/\varepsilon^2} a(sign\ Y_s)ds + \varepsilon \int_0^{t/\varepsilon^2} \sigma(sign\ Y_s)d\tilde{W}(s) + \varepsilon w_1(t), \tag{4.4}$$

with the standard Wiener process $\tilde{W}(s) = \varepsilon^{-1} W(\varepsilon^2 s)$.

Further,

$$E\varepsilon^2 \int_0^{t/\varepsilon^2} \sigma(sign\ Y_s)d\tilde{W}(s) \cdot (\int_0^{t/\varepsilon^2} \sigma(sign\ Y_s)d\tilde{W}(s))^T$$

$$= \varepsilon^2 E(\int_0^{t/\varepsilon^2} \sigma(sign\ Y_s)\cdot \sigma^T(sign\ Y_s)ds) \to [\sigma^-(\sigma^-)^T p^- + \sigma^+(\sigma^+)^T p^+]t := \bar{\sigma}\bar{\sigma}^T t, \tag{4.5}$$

where indeed p^- and p^+ are from (2.7) and $\bar{\sigma}$ is a solution of the matrix equation with respect to $\bar{\sigma}$

$$\bar{\sigma}\bar{\sigma}^T = \sigma^-(\sigma^-)^T p^- + \sigma^+(\sigma^+)^T p^+.$$

Here one has a convergence in probability locally uniformly with respect to t. Hence, in fact, we get a weak convergence to the Gaussian process described in the theorem. This proves the assertion.

4.2. General problem. Now let us proceed to a more general case. Consider a d-dimensional system of stochastic differential equations

$$dX = f(X)dt + \sigma(X)dW(t), \tag{4.6}$$

where X and f are d-dimensional vectors, W is a k-dimensional standard Wiener process, σ is a $d \times k$ matrix.

Let $S:\ \varphi(x) = 0$ be a smooth surface of discontinuity for the drift and diffusion coefficients of the system (4.6) and the domain G^+ is chosen so that the vector $\nabla\varphi(x)$ at the point $x \in S$ directed to G^+. We suppose the coefficients $f(x)$ and $\sigma(x)$ together with their first derivatives with respect to x to be continuous functions in G^- and G^+ up to S. In addition we suppose that at any point $x \in S$ the following conditions are fulfilled:

$$(\sigma^-)^\intercal \nabla\varphi = 0,\ (\sigma^+)^\intercal \nabla\varphi = 0, \tag{4.7}$$

$$(L\varphi)^- > 0,\ (L\varphi)^+ < 0, \tag{4.8}$$

where

$$L\varphi(x) = \sum_{i=1}^{d} \frac{\partial\varphi}{\partial x_i}(x) f_i(x) + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2\varphi}{\partial x_i \partial x_j}(x) \sum_{m=1}^{k} \sigma_{im}(x)\sigma_{jm}(x). \tag{4.9}$$

The condition (4.7) means that the orthogonal to the surface S component of diffusion degenerates with approaching S from both sides. The conditions (4.7) and (4.8) ensure the impossibility of leaving the surface S. Because of the degeneracy (4.7), the problem of a "natural" definition of solutions on S arises. In connection with this problem we introduce the auxiliary system with additional nondegenerate small noise

$$dX_t^\varepsilon = f(X_t^\varepsilon)dt + \sigma(X_t^\varepsilon)dW(t) + \varepsilon dw(t), \tag{4.10}$$

where w is a d-dimensional standard Wiener process independent on W.

To obtain the limit stochastic process on S, we suppose that the surface S can be expressed by an equation solved with respect to the last component in a neighborhood of a considered point. Let us represent the system (4.10) in the form

$$dX_t^\varepsilon = a(X_t^\varepsilon, Y_t^\varepsilon)dt + \alpha(X_t^\varepsilon, Y_t^\varepsilon)dW(t) + \varepsilon dw(t)$$

$$dY_t^\varepsilon = b(X_t^\varepsilon, Y_t^\varepsilon)dt + \beta(X_t^\varepsilon, Y_t^\varepsilon)dW(t) + \varepsilon dv(t). \tag{4.11}$$

Here $a,\ X$ are $(d-1)$-dimensional vectors, $b,\ Y$ are scalars, α is $(d-1)\times k$-matrix and β is $1\times k$-matrix, W is a k-dimensional, w is a $(d-1)$-dimensional and v is a scalar standard independent Wiener processes. Let S be expressed by the equation $Y = \psi(X)$ in a neighborhood of the considered point $(x, y) = (x_1, ..., x_{d-1}, y)$. Because

$$(\sigma^\pm)^\intercal = \begin{bmatrix} \alpha_{11}^\pm & ... & \alpha_{d-11}^\pm & \beta_1^\pm \\ & & & \\ \alpha_{1k}^\pm & ... & \alpha_{d-1k}^\pm & \beta_k^\pm \end{bmatrix},\ \nabla\varphi = \begin{bmatrix} -\partial\psi/\partial x_1 \\ ... \\ -\partial\psi/\partial x_{d-1} \\ 1 \end{bmatrix}, \tag{4.12}$$

we get from (4.7)

$$(\sigma^{\pm})^{\intercal}\nabla\varphi = -\sum_{j=1}^{d-1}\alpha_{jm}^{\pm}\frac{\partial\psi}{\partial X_j} + \beta_m^{\pm} = 0,\ m = 1, ..., k. \tag{4.13}$$

Further,

$$L\varphi = -\sum_{j=1}^{d-1} a_j\frac{\partial\psi}{\partial X_j} + b - \frac{1}{2}\sum_{i,j=1}^{d-1}\frac{\partial^2\psi}{\partial X_i\partial X_j}\sum_{m=1}^{k}\alpha_{im}\alpha_{jm}\ . \tag{4.14}$$

In the new variables

$$\tilde{X} = X,\ \tilde{Y} = Y - \psi(X)$$

we obtain (see (4.13), (4.14))

$$d\tilde{X}_t^{\varepsilon} = a(\tilde{X}_t^{\varepsilon}, \tilde{Y}_t^{\varepsilon} + \psi(\tilde{X}_t^{\varepsilon}))dt + \alpha(\tilde{X}_t^{\varepsilon}, \tilde{Y}_t^{\varepsilon} + \psi(\tilde{X}_t^{\varepsilon}))dW(t) + \varepsilon dw(t) \tag{4.15}$$

$$d\tilde{Y}_t^{\varepsilon} = L\varphi(\tilde{X}_t^{\varepsilon}, \tilde{Y}_t^{\varepsilon} + \psi(\tilde{X}_t^{\varepsilon}))dt$$

$$+(\nabla\varphi)^{\intercal}\sigma(\tilde{X}_t^{\varepsilon}, \tilde{Y}_t^{\varepsilon} + \psi(\tilde{X}_t^{\varepsilon}))dW(t) - \varepsilon\sum_{j=1}^{d-1}\frac{\partial\psi}{\partial X_j}(\tilde{X}_t^{\varepsilon})dw_j(t) + \varepsilon dv(t). \tag{4.16}$$

The surface of discontinuity for the system (4.15)-(4.16) is expressed by the equation $\tilde{Y} = 0$ in a neighborhood of the origin. Due to (4.13) the vector-row $(\nabla\varphi)^{\intercal}\sigma(\tilde{X}, \tilde{Y} + \psi(\tilde{X}))$ tends to zero if $\tilde{Y}$ tends to zero. Besides, the drift coefficient $L\varphi(\tilde{X}, \tilde{Y} + \psi(\tilde{X}))$ is positive for $\tilde{Y} < 0$ and negative for $\tilde{Y} > 0$ thanks to (4.8). Therefore the system (4.15)-(4.16) is close to the model system (4.1) and one can use the obtained rule for the sliding stochastic mode on the plane $\tilde{Y} = 0$. We get for

$$\breve{X}(t) := \lim_{\varepsilon\to 0}\tilde{X}_t^{\varepsilon}$$

the following stochastic differential equation

$$d\breve{X} = \breve{a}(\breve{X})dt + \breve{\sigma}(\breve{X})d\breve{w}(t), \tag{4.17}$$

where $\breve{w}$ is a $(d-1)$-dimensional standard Wiener process,

$$\breve{a}(\breve{X}) = a^{-}(\breve{X}, \psi(\breve{X}))p^{-} + a^{+}(\breve{X}, \psi(\breve{X}))p^{+}, \tag{4.18}$$

$$\breve{\sigma}(\breve{X})(\breve{\sigma}(\breve{X}))^{\intercal} = \alpha^{-}(\alpha^{-})^{\intercal}p^{-} + \alpha^{+}(\alpha^{+})^{\intercal}p^{+}, \tag{4.19}$$

with

$$\alpha^{\pm} = \alpha^{\pm}(\breve{X}, \psi(\breve{X}))$$

and

$$p^{-}(\breve{X}) = -\frac{(L\varphi)^{+}}{(L\varphi)^{-} - (L\varphi)^{+}}\ ,\ p^{+}(\breve{X}) = \frac{(L\varphi)^{-}}{(L\varphi)^{-} - (L\varphi)^{+}}\ , \tag{4.20}$$

where the arguments by $L\varphi$ are $(\tilde{X}, \tilde{Y} + \psi(\tilde{X}))$ under $\tilde{X} = \breve{X},\ \tilde{Y} = 0$, i.e., $(\breve{X}, \psi(\breve{X}))$.

Let

$$\bar{X}(t) := \lim_{\varepsilon\to 0} X_t^{\varepsilon},\ \bar{Y}(t) := \lim_{\varepsilon\to 0} Y_t^{\varepsilon}.$$

Due to the change of variables, $\bar{X} = \breve{X} + x$, $\bar{Y} = \psi(\bar{X})$, and we have

$$d\bar{X} = \bar{a}(\bar{X}, \bar{Y})dt + \bar{\alpha}(\bar{X}, \bar{Y})d\breve{w}(t), \tag{4.21}$$

where

$$\bar{a}(\bar{X}, \bar{Y}) = a^-(\bar{X}, \bar{Y})p^- + a^+(\bar{X}, \bar{Y})p^+ \tag{4.22}$$

$$\bar{\alpha}(\bar{X}, \bar{Y})\bar{\alpha}^\intercal(\bar{X}, \bar{Y}) = \alpha^-(\alpha^-)^\intercal p^- + \alpha^+(\alpha^+)^\intercal p^+, \tag{4.23}$$

with $\alpha^\pm$, $p^\pm$ depending on $\bar{X}, \bar{Y} = \psi(\bar{X})$.

Evaluate now $d\bar{Y} = d\psi(\bar{X})$. Thanks to (4.21)-(4.23) we get

$$\begin{aligned} d\bar{Y} &= \sum_{i=1}^{d-1} \frac{\partial\psi}{\partial X_i} \cdot (\bar{a}_i dt + \sum_{j=1}^{d-1} \bar{\alpha}_{ij} d\breve{w}_j(t)) + \frac{1}{2}\sum_{i,j=1}^{d-1} \frac{\partial^2\psi}{\partial X_i \partial X_j} \sum_{m=1}^{d-1} \bar{\alpha}_{im}\bar{\alpha}_{jm} dt \\ &= p^-(\sum_{i=1}^{d-1} \frac{\partial\psi}{\partial X_i} \cdot a_i^- + \frac{1}{2}\sum_{i,j=1}^{d-1} \frac{\partial^2\psi}{\partial X_i \partial X_j} \sum_{m=1}^{k} \alpha_{im}^- \alpha_{jm}^-)dt \\ &+ p^+(\sum_{i=1}^{d-1} \frac{\partial\psi}{\partial X_i} \cdot a_i^+ + \frac{1}{2}\sum_{i,j=1}^{d-1} \frac{\partial^2\psi}{\partial X_i \partial X_j} \sum_{m=1}^{k} \alpha_{im}^+ \alpha_{jm}^+)dt + \sum_{i=1}^{d-1} \frac{\partial\psi}{\partial X_i} \sum_{j=1}^{d-1} \bar{\alpha}_{ij} d\breve{w}_j(t), \end{aligned} \tag{4.24}$$

where all the functions have $\bar{X}, \bar{Y}$ as their arguments.

Due to (4.14)

$$\sum_{i=1}^{d-1} \frac{\partial\psi}{\partial X_i} \cdot a_i^\pm + \frac{1}{2}\sum_{i,j=1}^{d-1} \frac{\partial^2\psi}{\partial X_i \partial X_j} \sum_{m=1}^{k} \alpha_{im}^\pm \alpha_{jm}^\pm = b^\pm - (L\varphi)^\pm, \tag{4.25}$$

and due to (4.20)

$$(L\varphi)^- p^- + (L\varphi)^+ p^+ = 0. \tag{4.26}$$

The relations (4.24)-(4.26) give

$$d\bar{Y} = \bar{b}(\bar{X}, \bar{Y})dt + \sum_{i=1}^{d-1} \frac{\partial\psi}{\partial X_i}(\bar{X}) \sum_{j=1}^{d-1} \bar{\alpha}_{ij}(\bar{X}, \bar{Y}) d\breve{w}_j(t), \tag{4.27}$$

where

$$\bar{b}(\bar{X}, \bar{Y}) = b^-(\bar{X}, \bar{Y})p^-(\bar{X}, \bar{Y}) + b^+(\bar{X}, \bar{Y})p^+(\bar{X}, \bar{Y}). \tag{4.28}$$

The $d \times d$-dimensional diffusion matrix of the system (4.21), (4.27) is equal to $\Gamma\Gamma^\intercal$, where Γ is $d \times (d-1)$-dimensional:

$$\Gamma = \begin{bmatrix} \bar{\alpha}_{11} & \dots & \bar{\alpha}_{1d-1} \\ \dots & \dots & \dots \\ \bar{\alpha}_{d-11} & \dots & \bar{\alpha}_{d-1d-1} \\ \sum_{i=1}^{d-1} \frac{\partial\psi}{\partial X_i}\bar{\alpha}_{i1} & \dots & \sum_{i=1}^{d-1} \frac{\partial\psi}{\partial X_i}\bar{\alpha}_{id-1} \end{bmatrix}.$$

Using (4.23) and (4.13) one can prove that

$$\Gamma\Gamma^\intercal = \sigma^-(\sigma^-)^\intercal p^- + \sigma^+(\sigma^+)^\intercal p^+,$$

where $\sigma^\pm$ is the $d \times k$-dimensional matrix (see (4.12)).

Introduce a $d \times d$-dimensional matrix $\bar{\sigma}$ which satisfies the equation

$$\bar{\sigma}\bar{\sigma}^\intercal = \sigma^-(\sigma^-)^\intercal p^- + \sigma^+(\sigma^+)^\intercal p^+. \tag{4.29}$$

Then the system

$$d\bar{X}_i = \bar{a}_i(\bar{X}, \bar{Y})dt + \sum_{j=1}^{d} \bar{\sigma}_{ij}(\bar{X}, \bar{Y})d\bar{w}_j(t),\ i = 1, ..., d-1,$$

$$d\bar{Y} = \bar{b}(\bar{X}, \bar{Y})dt + \sum_{j=1}^{d} \bar{\sigma}_{dj}(\bar{X}, \bar{Y})d\bar{w}_j(t),$$

where $\bar{w}(t)$ is a d-dimensional standard Wiener process, gives the same diffusion law as the system (4.21), (4.27).

Now we can return to the original systems (4.6) and (4.10) and state the following assertion.

Theorem 4.2. *Suppose the conditions* (4.7), (4.8) *to be satisfied. Let* X_t^ε *be a solution of the system* (4.10) *starting from a point* $x \in S,\ 0 \le t \le \Delta t$, where $\Delta t > 0$ *is sufficiently small. Then there exists the limit*

$$\lim_{\varepsilon \to 0} X_t^\varepsilon := \bar{X}(t),$$

where $\bar{X}(0) = x$ *and* $\bar{X}(t) \in S,\ 0 \le t \le \Delta t$.

The process $\bar{X}(t)$ *is governed by the following system of stochastic differential equations*

$$d\bar{X} = \bar{f}(\bar{X})dt + \bar{\sigma}(\bar{X})d\bar{W}(t), \tag{4.30}$$

where $\bar{W}$ *is a d-dimensional standard Wiener process and*

$$\bar{f}(x) = f^-(x)p^-(x) + f^+(x)p^+(x),\ x \in S, \tag{4.31}$$

$$\bar{\sigma}(x)\bar{\sigma}^\intercal(x) = \sigma^-(x)(\sigma^-(x))^\intercal p^-(x) + \sigma^+(x)(\sigma^+(x))^\intercal p^+(x),\ x \in S, \tag{4.32}$$

$$p^-(x) = -\frac{(L\varphi(x))^+}{(L\varphi(x))^- - (L\varphi(x))^+},\ p^+(x) = \frac{(L\varphi(x))^-}{(L\varphi(x))^- - (L\varphi(x))^+},\ x \in S. \tag{4.33}$$

Remark 4.1. It is evident that the vector $\bar{f}$ in the system (2.29) is tangent to S and therefore the $(d-1)$-dimensional manifold is invariant for a system of the form (2.29) in whatever way (of course, in a sufficiently smooth manner) it is continued (we observe that $\bar{f}$ is determined on S only).

Analogously, in whatever way to continue the coefficients of the system (4.30) the manifold S remains invariant.

Let us check this fact directly using the Stroock-Varadhan support theorem [2]. In accord with the theorem one has to write the Ito system (4.30) (of course, with continued in a smooth manner coefficients) in the Stratonovich form and verify that the drift- and diffusion-vectors of the latter system are tangent to S.

Rewrite the system (4.30)

$$d\bar{X} = \bar{f}(\bar{X})dt + \sum_{r=1}^{d} \bar{\sigma}^r(\bar{X})d\bar{w}_r(t), \tag{4.34}$$

where $\bar{\sigma}^r$ is the r-th column of the matrix $\bar{\sigma}(\bar{X})$. We consider the matrix $\bar{\sigma}$ which satisfies (4.29) to be symmetric for simplicity, i.e., $\bar{\sigma} = \bar{\sigma}^\intercal$.

The Stratonovich system corresponding to (4.34) has the form

$$d\bar{X} = (\bar{f}(\bar{X}) - \frac{1}{2}\sum_{r=1}^{d} \frac{\partial \bar{\sigma}^r}{\partial x}(\bar{X})\bar{\sigma}^r(\bar{X}))dt + \sum_{r=1}^{d} \bar{\sigma}^r(\bar{X}) * d\bar{w}_r(t), \tag{4.35}$$

where

$$\frac{\partial \bar{\sigma}^r}{\partial x} = \begin{bmatrix} \frac{\partial \bar{\sigma}_1^r}{\partial x_1} & \cdots & \frac{\partial \bar{\sigma}_1^r}{\partial x_d} \\ \cdots & \cdots & \cdots \\ \frac{\partial \bar{\sigma}_d^r}{\partial x_1} & \cdots & \frac{\partial \bar{\sigma}_d^r}{\partial x_d} \end{bmatrix}.$$

Let us check that on S

$$(\bar{\sigma}^r, \nabla\varphi) = 0, \ r = 1, ..., d. \tag{4.36}$$

The relations (4.36) are equivalent to

$$\bar{\sigma}^\intercal \nabla\varphi = 0.$$

We have

$$(\bar{\sigma}^\intercal\nabla\varphi, \bar{\sigma}^\intercal\nabla\varphi) = (\bar{\sigma}\bar{\sigma}^\intercal\nabla\varphi, \nabla\varphi)$$

$$= ((p^-\sigma^-(\sigma^-)^\intercal\nabla\varphi + p^+\sigma^+(\sigma^+)^\intercal\nabla\varphi), \nabla\varphi) = 0,$$

as $(\sigma^\pm)^\intercal\nabla\varphi = 0$ on S (see the condition (4.7)).

So, $\bar{\sigma}^\intercal\nabla\varphi = 0$ and, consequently, (4.36) is proved.

Let us evaluate

$$-\frac{1}{2}(\sum_{r=1}^{d} \frac{\partial \bar{\sigma}^r}{\partial x}\bar{\sigma}^r, \nabla\varphi) = -\frac{1}{2}\sum_{r=1}^{d}(\bar{\sigma}^r, \begin{bmatrix} \frac{\partial \bar{\sigma}_1^r}{\partial x_1}\frac{\partial\varphi}{\partial x_1} + ... + \frac{\partial \bar{\sigma}_d^r}{\partial x_1}\frac{\partial\varphi}{\partial x_d} \\ \cdots\cdots\cdots \\ \frac{\partial \bar{\sigma}_1^r}{\partial x_d}\frac{\partial\varphi}{\partial x_1} + ... + \frac{\partial \bar{\sigma}_d^r}{\partial x_d}\frac{\partial\varphi}{\partial x_d} \end{bmatrix})$$

$$= -\frac{1}{2}\sum_{r=1}^{d}(\bar{\sigma}^r, \begin{bmatrix} \frac{\partial}{\partial x_1}(\bar{\sigma}_1^r\frac{\partial\varphi}{\partial x_1} + ... + \bar{\sigma}_d^r\frac{\partial\varphi}{\partial x_d}) \\ \cdots\cdots\cdots \\ \frac{\partial}{\partial x_d}(\bar{\sigma}_1^r\frac{\partial\varphi}{\partial x_1} + ... + \bar{\sigma}_d^r\frac{\partial\varphi}{\partial x_d}) \end{bmatrix})$$

$$+\frac{1}{2}\sum_{r=1}^{d}(\bar{\sigma}^r, \begin{bmatrix} \bar{\sigma}_1^r\frac{\partial^2\varphi}{\partial x_1^2} + ... + \bar{\sigma}_d^r\frac{\partial^2\varphi}{\partial x_1\partial x_d} \\ \cdots\cdots\cdots \\ \bar{\sigma}_1^r\frac{\partial^2\varphi}{\partial x_d\partial x_1} + ... + \bar{\sigma}_d^r\frac{\partial^2\varphi}{\partial x_d^2} \end{bmatrix})$$

$$-\frac{1}{2}\sum_{r=1}^{d}(\bar{\sigma}^r, \frac{\partial}{\partial x}(\bar{\sigma}^r, \nabla\varphi)) + \frac{1}{2}\mathrm{tr}\bar{\sigma}\bar{\sigma}^\intercal\left\{\frac{\partial^2\varphi}{\partial x_i\partial x_j}\right\}. \tag{4.37}$$

The relation (4.36) asserts that the vector $\bar{\sigma}^r$ on S, $r = 1, ..., d$, is tangent to S. The sum $\sum_{r=1}^{d}(\bar{\sigma}^r, \frac{\partial}{\partial x}(\bar{\sigma}^r, \nabla\varphi))$ on S is none other than the derivative of the function

$(\bar{\sigma}^r, \nabla\varphi)$ in the direction of $\bar{\sigma}^r$, i.e., in the direction tangent to S. But on S this function is equal to zero (see (4.36)). Hence the first summand in the right hand side of (4.37) is equal to zero on S.

Consider the second summand in (4.37). From (4.9) we have

$$\frac{1}{2}\mathrm{tr}\sigma^{\pm}(\sigma^{\pm})^{\mathsf{T}}\left\{\frac{\partial^2\varphi}{\partial x_i \partial x_j}\right\} = (L\varphi)^{\pm} - (f^{\pm}, \nabla\varphi).$$

Therefore (see (4.32), (4.26), (4.31))

$$\frac{1}{2}\mathrm{tr}\bar{\sigma}\bar{\sigma}^{\mathsf{T}}\left\{\frac{\partial^2\varphi}{\partial x_i \partial x_j}\right\} = p^- \cdot \frac{1}{2}\mathrm{tr}\sigma^-(\sigma^-)^{\mathsf{T}}\left\{\frac{\partial^2\varphi}{\partial x_i \partial x_j}\right\} + p^+ \cdot \frac{1}{2}\mathrm{tr}\sigma^+(\sigma^+)^{\mathsf{T}}\left\{\frac{\partial^2\varphi}{\partial x_i \partial x_j}\right\}$$

$$= p^- \cdot ((L\varphi)^- - (f^-, \nabla\varphi)) + p^+ \cdot ((L\varphi)^+ - (f^+, \nabla\varphi))$$

$$= -p^- \cdot (f^-, \nabla\varphi) + p^+ \cdot (f^+, \nabla\varphi) = -(\bar{f}, \nabla\varphi).$$

Consequently,

$$(\bar{f} - \frac{1}{2}\sum_{r=1}^{d}\frac{\partial\bar{\sigma}^r}{\partial x}\bar{\sigma}^r, \nabla\varphi) = 0,$$

i.e., the drift in the system (4.35) is tangent to S as well. Thus, the invariance of S for the system (4.30) is verified.

5. Sliding mode on surfaces which dimension is less than $d-1$

In this section we treat a problem connected with determining sliding mode on a $(d-2)$-dimensional surface which is the intersection of two $(d-1)$-dimensional sliding surfaces. We restrict ourselves to a model problem in the case $d = 3$.

Consider the following three-dimensional system

$$\frac{dx}{dt} = a(y,z), \ \frac{dy}{dt} = b(y,z), \ \frac{dz}{dt} = c(y,z) \tag{5.1}$$

with

$$a(y,z) = a_1, \ b(y,z) = b_1, \ c(y,z) = c_1 \text{ if } y > 0, \ z > 0,$$

$$a(y,z) = a_2, \ b(y,z) = b_2, \ c(y,z) = c_2 \text{ if } y < 0, \ z > 0,$$

$$a(y,z) = a_3, \ b(y,z) = b_3, \ c(y,z) = c_3 \text{ if } y < 0, \ z < 0,$$

$$a(y,z) = a_4, \ b(y,z) = b_4, \ c(y,z) = c_4 \text{ if } y > 0, \ z < 0,$$

where $a_i, \ b_i, \ c_i, \ i = 1, ..., 4$, are constants, i.e., the right-hand sides of the system (5.1) depend in fact on *sign* y and *sign* z only.

In addition we suppose that

$$b_1 < 0, \ b_2 > 0, \ b_3 > 0, \ b_4 < 0,$$

$$c_1 < 0, \ c_2 < 0, \ c_3 > 0, \ c_4 > 0. \tag{5.2}$$

Thus, the surfaces of discontinuity of the right-hand sides of the system (5.1) are the planes $y = 0$ and $z = 0$. They divide the space in four parts. The conditions (5.2) ensure that any trajectory reaches one of these planes for a finite time. The law of motion on the planes can be obtained according to the results of Section 2. It is not difficult to see that any motion reaches the line $y = 0$, $z = 0$ (axis x) for a finite time as well and our problem is to find the law of motion on this sliding line.

Introduce the system with small noise

$$dX_t^\varepsilon = a(Y_t^\varepsilon, Z_t^\varepsilon)dt + \varepsilon dw_1(t), \tag{5.3}$$

$$dY_t^\varepsilon = b(Y_t^\varepsilon, Z_t^\varepsilon)dt + \varepsilon dw_2(t),\ dZ_t^\varepsilon = c(Y_t^\varepsilon, Z_t^\varepsilon)dt + \varepsilon dw_3(t) \tag{5.4}$$

and the process (Y_s, Z_s) :

$$Y_s = \varepsilon^{-2} Y_{\varepsilon^2 s}^\varepsilon,\ Z_s = \varepsilon^{-2} Z_{\varepsilon^2 s}^\varepsilon. \tag{5.5}$$

Because of (5.2), it is not difficult to justify that the law of the process (Y_s, Z_s) does not depend on ε and that it satisfies the system

$$dY = b(Y, Z)ds + dW_2(s),\ dZ = c(Y, Z)ds + dW_3(s), \tag{5.6}$$

where $W_2(s) = \varepsilon^{-1} w_2(\varepsilon^2 s)$, $W_3(s) = \varepsilon^{-1} w_3(\varepsilon^2 s)$.

Due to the conditions (5.2) the Markov process defined by the system (5.6) is ergodic. Let μ be its invariant measure. Let $(X_t^\varepsilon, Y_t^\varepsilon, Z_t^\varepsilon)$ be the solution of the system (5.3)-(5.4) starting from the origin. Analogously to (2.11) we get

$$\lim_{\varepsilon \to 0} X_t^\varepsilon = \sum_{i=1}^4 a_i p_i \cdot t := \bar{a} t,\ \lim_{\varepsilon \to 0} Y_t^\varepsilon = 0,\ \lim_{\varepsilon \to 0} Z_t^\varepsilon = 0,$$

where p_i is μ-measure of the corresponding quadrant of the plane $X = 0$. Obtain some relation for p_i. Let $Y(0) = \xi$, $Z(0) = \eta$, where (ξ, η) is a random vector with the distribution law μ. We have from (5.6)

$$Y(s) = \xi + b_1 \int_0^s \chi_{Y(\theta)>0, Z(\theta)>0} d\theta + \ldots + b_4 \int_0^s \chi_{Y(\theta)>0, Z(\theta)<0} d\theta \ + W_2(s).$$

Because $EY(s) = E\xi$, $E\chi_{Y(\theta)>0, Z(\theta)>0} = \mu(Y > 0, Z > 0) = p_1, \ldots,$ $E\chi_{Y(\theta)>0, Z(\theta)<0} = \mu(Y > 0, Z < 0) = p_4$, we get

$$b_1 p_1 + b_2 p_2 + b_3 p_3 + b_4 p_4 = 0.$$

Similarly

$$c_1 p_1 + c_2 p_2 + c_3 p_3 + c_4 p_4 = 0.$$

Together with $p_1 + p_2 + p_3 + p_4 = 1$ we have three relations with respect to four desired probabilities. Unfortunately, there is no additional relation in general case. However, for instance under some kind of symmetry, it is possible to find these probabilities or $\bar{a}$. A short example: let $b_2 = -b_1$, $c_2 = c_1$, $b_4 = -b_3$, $c_4 = c_3$. Then it is clear that $p_1 = p_2$ and $p_3 = p_4$. As a result we obtain $p_1 = p_2 = \dfrac{c_3}{2(c_3 - c_1)}$, $p_3 = p_4 = \dfrac{-c_1}{2(c_3 - c_1)}$, and $\bar{a}$ can be found explicitly. Another example. It is not difficult to prove that the following three four-dimensional vectors: $\tilde{b} := (b_1, b_2, b_3, b_4)^\intercal$, $\tilde{c} := (c_1, c_2, c_3, c_4)^\intercal$, $\tilde{p} := (1, 1, 1, 1)^\intercal$ are linearly independent provided (5.2). Let $\tilde{a} := (a_1, a_2, a_3, a_4)^\intercal$ depend linearly on $\tilde{b}$, $\tilde{c}$, $\tilde{p}$: $\tilde{a} = \beta \tilde{b} + \gamma \tilde{c} + \alpha \tilde{p}$. Then $\bar{a} = \alpha$.

References

[1] A.F. Filippov. Differential equations with discontinuous right-hand side. Nauka, Moscow, 1985.

[2] N.Ikeda, S.Watanabe. Stochastic differential equations and diffusion processes. North-Holland Publ. Comp., Amsterdam, 1981.

[3] N.V. Krylov. Ito's stochastic integral equations. Teor. Verojatnost. i Primenen. 14(1969), 340-348.

[4] N.V. Krylov. The selection of a Markov process from a Markov system of processes, and the construction of quasidiffusion processes. Izv. Akad. Nauk SSSR Ser. Mat. 37(1973), 691-708.

[5] D.W. Stroock, S.R.S. Varadhan. Multidimensional diffusion processes. Springer. 1979.

[6] A.Yu. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. Mat. Sb. (N.S.) 111(153) (1980), no.3, 434-452, 480.

[7] A.Yu. Veretennikov. Bounds for the mixing rate in the theory of stochastic equations, Theory Probab. Appl. 32(2) (1987), 273-281.

[8] A.Yu. Veretennikov. On polynomial mixing bounds for stochastic differential equations. Stoch. Processes Appl. (1997), 70, 115-127.