Taylor & Francis
Taylor & Francis Group

Higher-order sliding modes, differentiation and output-feedback control

ARIE LEVANT†

Being a motion on a discontinuity set of a dynamic system, sliding mode is used to keep accurately a given constraint and features theoretically-infinite-frequency switching. Standard sliding modes provide for finite-time convergence, precise keeping of the constraint and robustness with respect to internal and external disturbances. Yet the relative degree of the constraint has to be 1 and a dangerous chattering effect is possible. Higher-order sliding modes preserve or generalize the main properties of the standard sliding mode and remove the above restrictions. r-Sliding mode realization provides for up to the rth order of sliding precision with respect to the sampling interval compared with the first order of the standard sliding mode. Such controllers require higher-order real-time derivatives of the outputs to be available. The lacking information is achieved by means of proposed arbitrary-order robust exact differentiators with finite-time convergence. These differentiators feature optimal asymptotics with respect to input noises and can be used for numerical differentiation as well. The resulting controllers provide for the full output-feedback real-time control of any output variable of an uncertain dynamic system, if its relative degree is known and constant. The theoretical results are confirmed by computer simulation.

1. Introduction

Control under uncertainty condition is one of the main topics of the modern control theory. In spite of extensive and successful development of robust adaptive control and backstepping technique (Landau *et al.* 1998, Kokotovic and Arcak 2001) the sliding-mode control approach stays probably the main choice when one needs to deal with non-parametric uncertainties and unmodelled dynamics. That approach is based on keeping exactly a properly chosen constraint by means of high-frequency control switching. It exploits the main features of the sliding mode: its insensitivity to external and internal disturbances, ultimate accuracy and finite-time transient. However, the standard sliding-mode usage is bounded by some restrictions. The constraint being given by an equality of an output variable σ to zero, the standard sliding mode may be implemented only if the relative degree of σ is 1. In other words, control has to appear explicitly already in the first total derivative $\dot{\sigma}$. Also, high frequency control switching leads to the so-called chattering effect which is exhibited by high frequency vibration of the controlled plant and can be dangerous in applications.

A number of methods were proposed to overcome these difficulties. In particular, high-gain control with saturation approximates the sign-function and diminishes the chattering, while on-line estimation of the so-called equivalent control (Utkin 1992) is used to reduce the discontinuous-control component (Slotine and Li 1991), the sliding-sector method (Furuta and Pan 2000) is suitable to control disturbed linear time-

invariant systems. Yet, the sliding-mode order approach (Levantovsky 1985, Levant 1993) seems to be the most comprehensive, for it allows to remove all the above restrictions, while preserving the main sliding-mode features and improving its accuracy. Independently developed dynamical (Sira-Ramírez 1993, Rios-Bolívar *et al.* 1997) and terminal (Man *et al.* 1994) sliding modes are closely related to this approach.

Suppose that $\sigma \equiv 0$ is kept by a discontinuous dynamic system. While successively differentiating σ along trajectories, a discontinuity will be encountered sooner or later in the general case. Thus, sliding modes $\sigma \equiv 0$ may be classified by the number r of the first successive total derivative $\sigma^{(r)}$ which is not a continuous function of the state space variables or does not exist due to some reason like trajectory non-uniqueness. That number is called sliding order (Levant 1993, Fridman and Levant 1996, Bartolini *et al.* 1999 a). The standard sliding mode on which most variable structure systems (VSS) are based is of the first order ($\dot{\sigma}$ is discontinuous). While the standard modes feature finite time convergence, convergence to higher-order sliding modes (HOSM) may be asymptotic as well. While the standard sliding mode precision is proportional to the time interval between the measurements or to the switching delay, r-sliding mode realization may provide for up to the rth order of sliding precision with respect to the measurement interval (Levant 1993). Properly used, HOSM totally removes the chattering effect.

Trivial cases of asymptotically stable HOSM are often found in standard VSSs. For example there is an asymptotically stable 2-sliding mode with respect to the constraint $x = 0$ at the origin $x = \dot{x} = 0$ (at one point only) of a two-dimensional VSS keeping the constraint $x + \dot{x} = 0$ in a standard 1-sliding mode. Asymptotically stable or unstable HOSMs inevitably appear in VSSs

with fast actuators (Fridman 1990, Fridman and Levant 1996, 2002). Stable HOSM leads in that case to spontaneous disappearance of the chattering effect. Asymptotically stable or unstable sliding modes of any order are well known (Emelyanov *et al.* 1986, Elmali and Olgac 1992, Fridman and Levant 1996). Dynamic sliding modes (Sira-Ramírez 1993, Spurgeon and Lu 1997) produce asymptotically stable higher-order sliding modes and are to be specially mentioned here.

A family of finite-time convergent sliding mode controllers is based on so-called 'terminal sliding modes' (Man *et al.* 1994, Wu *et al.* 1998). Having been independently developed, the first version of these controllers is close to the so-called '2-sliding algorithm with a prescribed convergence law' (Emelyanov *et al.* 1986, Levant 1993). The latter version is intended actually to provide for arbitrary-order sliding mode with finite-time convergence. Unfortunately, the resulting closed-loop systems have unbounded right-hand sides, which prevents the very implementation of the Filippov theory. Thus, such a mode cannot be considered as HOSM. The corresponding control is formally bounded along each transient trajectory, but takes on infinite values in any vicinity of the steady state. In order to avoid infinite control values all trajectories are to start from a prescribed sector of the state space. The very definition and the existence of the solution require some special study here.

Arbitrary-order sliding controllers with finite-time convergence were only recently demonstrated (Levant 1998 b, 2001 a). The proofs of these results are for the first time published in the present paper. These controllers provide for full output control of uncertain single-input–single-output (SISO) weakly-minimum-phase dynamic systems with a known constant relative degree r. The control influence is a discontinuous function of the output and its $r - 1$ real-time-calculated successive derivatives. The controller parameters may be chosen in advance, so that only one parameter is to be adjusted in order to control any system with a given relative degree. No detailed mathematical model is needed. The system's relative degree being artificially increased, sliding control of arbitrary smoothness order can be achieved, completely removing the chattering effect. Since many mechanical systems have constant relative degree due, actually, to the Newton law, the application area for these controllers is very wide.

Any implementation of the above controllers requires real-time robust estimation of the higher-order total output derivatives. The popular high-gain observers (Dabroom and Khalil 1997) would destroy the exactness and finite-time-convergence features of the proposed controllers. The first-order robust exact differentiator (Levant 1998 a) can be used here, but its successive application is cumbersome and not effective.

The problem is solved by recently presented arbitrary-order robust exact finite-time-convergent differentiators (Levant 1999, 2001 a,b). The proposed lth-order differentiator allows real-time robust exact differentiation up to the order l, provided the next $(l + 1)$th input derivative is bounded. Its performance is proved to be asymptotically optimal in the presence of small Lebesgue-measurable input noises. This paper is the first regular publication of these differentiators and of the corresponding proofs. Their features allow broad implementation in the non-linear feedback control theory due to the separation principle (Atassi and Khalil 2000) triviality. They can be also successfully applied for numerical differentiation.

The rth-order sliding controller combined with the $(r - 1)$th-order differentiator produce an output-feedback universal controller for SISO processes with permanent relative degree (Levant 2002). Having been only recently obtained, the corresponding results are only briefly described in this paper, for the author intends to devote a special paper to this subject. The features of the proposed universal controllers and differentiators are illustrated by computer simulation.

2. Preliminaries: higher-order sliding modes

Let us recall first that according to the definition by Filippov (1988) any discontinuous differential equation $\dot{x} = v(x)$, where $x \in \mathbb{R}^n$ and v is a locally bounded measurable vector function, is replaced by an equivalent differential inclusion $\dot{x} \in V(x)$. In the simplest case, when v is continuous almost everywhere, $V(x)$ is the convex closure of the set of all possible limits of $v(y)$ as $y \to x$, while $\{y\}$ are continuity points of v. Any solution of the equation is defined as an absolutely continuous function $x(t)$, satisfying the differential inclusion almost everywhere.

Consider a smooth dynamic system $\dot{x} = v(x)$ with a smooth output function σ, and let the system be closed by some possibly-dynamical discontinuous feedback (figure 1). Then, provided that successive total time derivatives $\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}$ are continuous functions of the closed-system state space variables, and the *r-sliding point set*

$$\sigma = \dot{\sigma} = \ddot{\sigma} = \cdots = \sigma^{(r-1)} = 0 \qquad (1)$$

is non-empty and consists locally of Filippov trajectories, the motion on set (1) is called *r-sliding mode* (*rth-order sliding mode*, Levantovsky 1985, Levant 1993, Fridman and Levant 1996).

The additional condition of the Filippov velocity set V containing more than one vector may be imposed in order to exclude some trivial cases. It is natural to call the sliding order r *strict* if $\sigma^{(r)}$ is discontinuous or does
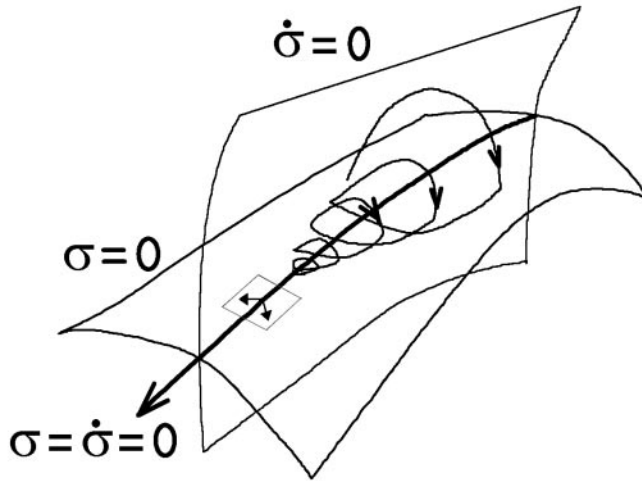
Figure 1.   2-sliding mode.

not exist in a vicinity of the r-sliding point set, but sliding-mode orders are mostly considered strict by default.

Hence, r-sliding modes are determined by equalities (1) which impose an r-dimensional condition on the state of the dynamic system. The sliding order characterizes the dynamics smoothness degree in some vicinity of the sliding mode.

Suppose that $\sigma, \dot{\sigma}, \ddot{\sigma}, \ldots, \sigma^{(r-1)}$ are differentiable functions of x and that

$$\text{rank} \, [\nabla \sigma, \nabla \dot{\sigma}, \ldots, \nabla \sigma^{(r-1)}] = r \qquad (2)$$

Equality (2) together with the requirement for the corresponding derivatives of σ to be differentiable functions of x is referred to as *r-sliding regularity condition*. If regularity condition (2) holds, then the r-sliding set is a differentiable manifold and $\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}$ may be supplemented up to new local coordinates.

Proposition 1 (Fridman and Levant 1996): *Let regularity condition* (2) *be fulfilled and r-sliding manifold* (1) *be non-empty. Then an r-sliding mode with respect to the constraint function σ exists if and only if the intersection of the Filippov vector-set field with the tangential space to manifold* (1) *is not empty for any r-sliding point.*

Proof: The intersection of the Filippov set of admissible velocities with the tangential space to the sliding manifold (1) induces a differential inclusion on this manifold (figure 1). This inclusion satisfies all the Filippov conditions for solution existence. Therefore manifold (1) is an integral one.   □

A sliding mode is called stable if the corresponding integral sliding set is stable. The above definitions are easily extended to include non-autonomous differential equations by introduction of the fictitious equation $\dot{t} = 1$. All the considerations are literally translated to the case of the closed-loop controlled system $\dot{x} = f(t, x, u)$, $u = U(t, x)$ with discontinuous U and smooth f, σ.

Real sliding: Up to this moment only ideal sliding modes were considered which keep $\sigma \equiv 0$. In reality, however, switching imperfections being present, ideal sliding cannot be attained. The simplest switching imperfection is discrete switching caused by discrete measurements. It was proved (Levant 1993) that the best possible sliding accuracy attainable with discrete switching in $\sigma^{(r)}$ is given by the relation $|\sigma| \sim \tau^r$, where $\tau > 0$ is the minimal switching time interval. Moreover, the relations $|\sigma^{(k)}| \sim \tau^{r-k}, k = 0, 1, \ldots, r$ are satisfied at the same time ($\sigma^{(0)} = \sigma$). Thus, in order to achieve the rth order of sliding precision in discrete realization, the sliding mode order in the continuous-time VSS has to be at least r. The standard sliding modes provide for the first-order real sliding only. The second order of real sliding was achieved by discrete switching modifications of the 2-sliding algorithms (Levant 1993, Bartolini *et al.* 1998) and by a special discrete switching algorithm (Su *et al.* 1994). Real sliding of higher orders is demonstrated in Fridman and Levant (1998) and Levant (1998 b, 2001 a).

In practice the final sliding accuracy is always achieved in finite time. When asymptotically stable modes are considered, however, it is not observable at any fixed moment, for the convergence time tends to infinity with the rise in accuracy. In the known cases the limit accuracy of the asymptotically stable modes can be shown to be of the first order only (Slotine and Li 1991). The above-mentioned highest precision is probably obtained only with finite-time-convergence sliding modes.

3.   The problem statement and its solutions with low relative degrees

Consider a dynamic system of the form

$$\dot{x} = a(t, x) + b(t, x)u, \quad \sigma = \sigma(t, x) \qquad (3)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}$, a, b and σ are smooth unknown functions, the dimension n is also unavailable. The relative degree r of the system is assumed to be constant and known. The task is to fulfil the constraint $\sigma(t, x) = 0$ in finite time and to keep it exactly by some feedback.

Extend system (3) by introduction of a fictitious variable $x_{n+1} = t$, $\dot{x}_{n+1} = 1$. Denote $a_e = (a, 1)^t$, $b_e = (b, 0)^t$, where the last component corresponds to x_{n+1}. The equality of the relative degree to r means that the Lie derivatives $L_{b_e}\sigma$, $L_{b_e}L_{a_e}\sigma, \ldots, L_{b_e}L_{a_e}^{r-2}$ equal zero identi-

cally in a vicinity of a given point and $L_{b_e}L_{a_e}^{r-1}\sigma$ is not zero at the point (Isidori 1989).

In a simplified way the equality of the relative degree to r means that u first appears explicitly only in the rth total derivative of σ. In that case regularity condition (2) is satisfied (Isidori 1989) and $(\partial/\partial u)\sigma^{(r)} \neq 0$ at the given point. The output σ satisfies an equation of the form

$$\sigma^{(r)} = h(t,x) + g(t,x)u \qquad (4)$$

It is easy to check that $g = L_{b_e}L_{a_e}^{r-1}\sigma = (\partial/\partial u)\sigma^{(r)}$, $h = L_{a_e}^r\sigma$. Obviously, h is the rth total time derivative of σ calculated with $u = 0$. In other words, unknown functions h and g may be defined using only input–output relations. The heavy uncertainty of the problem prevents immediate reduction of (3) to any standard form by means of standard approaches based on the knowledge of a, b and σ. Nevertheless, the very existence of standard form (4) is important here.

Proposition 2 (Fridman and Levant 1996): *Let system (3) have relative degree* r *with respect to the output function* σ *at some* r-*sliding point* (t_0, x_0). *Let, also,* $u = U(t,x)$, *the discontinuous function* U *taking on values from sets* $[K, \infty)$ *and* $(-\infty, -K]$ *on some sets of non-zero measure in any vicinity of each* r-*sliding point near point* (t_0, x_0). *Then this provides, with sufficiently large* K, *for the existence of* r-*sliding mode in some vicinity of the point* (t_0, x_0).

Proof: Proposition 2 is a straight-forward consequence of Proposition 1 and equation (4). □

The trivial controller $u = -K\mathrm{sign}\,\sigma$ satisfies Proposition 2. Usually, however, such a mode is not stable. The r-sliding mode motion is described by the equivalent control method (Utkin 1992), on the other hand, this dynamics coincides with the zero-dynamics (Isidori 1989) of the corresponding systems.

The problem is to find a discontinuous feedback $u = U(t,x)$ causing the appearance of a finite-time convergent r-sliding mode in (3). That new controller has to generalize the standard 1-sliding relay controller $u = -K\mathrm{sign}\,\sigma$. Hence, $g(t,y)$ and $h(t,y)$ in (4) are assumed to be bounded, $g > 0$. Thus, it is required that for some $K_m, K_M, C > 0$

$$0 < K_m \leq \frac{\partial}{\partial u}\sigma^{(r)} \leq K_M, \qquad |\sigma^{(r)}|_{u=0}| \leq C \qquad (5)$$

3.1. *Solutions of the problem with relative degrees* $r = 1, 2$

In case $r = 1$ calculation shows that

$$\dot\sigma(t,x,u) = \sigma'_t(t,x) + \sigma'_x(t,x)a(t,x) + \sigma'_x(t,x)b(t,x)u$$

and the problem is easily solved by the standard relay controller $u = -\alpha\,\mathrm{sign}\,\sigma$, with $a > C/K_m$. Here $h = \dot\sigma|_{u=0} = \sigma'_t + \sigma'_x a$ is globally bounded and $g = (\partial/\partial u)\dot\sigma = \sigma'_x b$. The first-order real-sliding accuracy with respect to the sampling interval is ensured.

Let $r = 2$. The following list includes only few most known controllers. The so-called twisting controller (Levantovsky 1985, Emelyanov *et al.* 1986, Levant 1993) and the convergence conditions are given by

$$u = -(r_1\mathrm{sign}\,\sigma + r_2\,\mathrm{sign}\,\dot\sigma), \qquad r_1 > r_2 > 0, \quad (6)$$

$$(r_1 + r_2)K_m - C > (r_1 - r_2)K_M + C,$$

$$(r_1 - r_2)K_m > C$$

A particular case of the controller with prescribed convergence law (Emelyanov *et al.* 1986, Levant 1993) is given by

$$u = -\alpha\,\mathrm{sign}\,(\dot\sigma + \lambda|\sigma|^{1/2}\,\mathrm{sign}\,\sigma), \qquad \alpha, \lambda > 0, \quad (7)$$

$$\alpha K_m - C > \lambda^2/2$$

Controller (7) is close to terminal sliding mode controllers (Man *et al.* 1994). The so-called sub-optimal controller (Bartolini *et al.* 1998, 1999 a, b) is given by

$$u = -r_1\,\mathrm{sign}\,(\sigma - \sigma^*/2) + r_2\,\mathrm{sign}\,\sigma^*, \qquad r_1 > r_2 > 0 \quad (8)$$

$$2[(r_1 + r_2)K_m - C] > (r_1 - r_2)K_M + C,$$

$$(r_1 - r_2)K_m > C \qquad (9)$$

where σ^* is the current value of σ detected at the closest time when $\dot\sigma$ was 0. The initial value of σ^* is 0. Any computer implementation of this controller requires successive measurements of $\dot\sigma$ or σ with some time step. Usually the detection of the moments when $\dot\sigma$ changes its sign is performed. The control value u depends actually on the history of $\dot\sigma$ and σ measurements, i.e. on $\dot\sigma(\cdot)$ and $\sigma(\cdot)$.

Theorem 1 (Levant 1993, Bartolini *et al.* 1998): *2-sliding controllers* (6)–(8) *provide for finite-time convergence of any trajectory of* (3) *and* (5) *to 2-sliding mode* $\sigma \equiv 0$. *The convergence time is a locally bounded function of the initial conditions.*

Let the measurements be carried out at times t_i with constant step $\tau > 0$, $\sigma_i = \sigma(t_i, x(t_i))$, $\Delta\sigma_i = \sigma_i - \sigma_{i-1}$, $t \in [t_i, t_{i+1})$. Substituting σ_i for σ, sign $\Delta\sigma_i$ for sign $\dot\sigma$, and sign $(\Delta\sigma_i - \lambda\tau|\sigma_i|^{1/2}\,\mathrm{sign}\,\sigma_i)$ for sign $(\dot\sigma - \lambda|\sigma|^{1/2}\,\mathrm{sign}\,\sigma)$ achieve discrete-sampling versions of the controllers.

Theorem 2 (Levant 1993, Bartolini *et al.* 1998): *Discrete-sampling versions of controllers* (6)–(8) *provide for*

the establishment of the inequalities $|\sigma| < \mu_0 \tau^2, |\dot\sigma| < \mu_1 \tau$ for some positive μ_0, μ_1.

All listed controllers may be used also with relative degree 1 in order to remove the chattering and improve the sliding accuracy. Indeed, let $u = \varphi(\sigma(\cdot), \dot\sigma(\cdot))$ be one of controllers (6)–(8), depending possibly on the previous measurements as in (8), then under certain natural conditions (Levant 1993, Bartolini *et al.* 1999 a) the controller $u = -\operatorname{sign}\sigma$ can be replaced by the controller $\dot u = -u$ with $|u| > 1, \dot u = \varphi(\sigma(\cdot), \dot\sigma(\cdot))$ with $|u| \le 1$.

Consider the general case. Let u be defined from the equality

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[P_{r-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)\sigma\right] = -\operatorname{sign}\left[P_{r-1}\left(\frac{\mathrm{d}t}{\mathrm{d}t}\right)\sigma\right],$$

where $P_{r-1}(\lambda) = \lambda^{(r-1)} + \alpha_1 \lambda^{(r-2)} + \cdots + \alpha_{r-1}$ is a stable polynomial, $\alpha_i \in \mathbb{R}$. In the case $r = 1$, $P_0 = 1$ achieve the standard 1-sliding mode. Let $r > 1$. The r-sliding mode exists here at the origin and is asymptotically stable. There is also a 1-sliding mode on the manifold $P_{r-1}(\mathrm{d}/\mathrm{d}t)\sigma = 0$. Trajectories transfer in finite time into the 1-sliding mode on the manifold $P_{r-1}(\mathrm{d}/\mathrm{d}t)\sigma = 0$ and then exponentially converge to the r-sliding mode. Dynamic-sliding-mode controllers (Sira-Ramírez 1993, Spurgeon and Lu 1997) are based on such modes. Unfortunately, due to the dependence on higher-order derivatives of σ the control is not bounded here even with small σ. Also the accuracy here is the same as of the 1-sliding mode.

4. Building an arbitrary-order sliding controller

Let p be any positive number, $p \ge r$. Denote

$$N_{1,r} = |\sigma|^{(r-1)/r}$$

$$N_{i,r} = (|\sigma|^{p/r} + |\dot\sigma|^{p/(r-1)} + \cdots + |\sigma^{(i-1)}|^{p/(r-i+1)})^{(r-i)/p},$$

$$i = 1, \ldots, r-1$$

$$N_{r-1,r} = (|\sigma|^{p/r} + \dot\sigma|^{(p/(r-1)} + \cdots + |\sigma^{(r-2)}|^{p/2})^{1/p}$$

$$\phi_{0,r} = \sigma$$

$$\phi_{1,r} = \dot\sigma + \beta_1 N_{1,r} \operatorname{sign}(\sigma)$$

$$\phi_{i,r} = \sigma^{(i)} + \beta_i N_{i,r} \operatorname{sign}(\phi_{i-1,r}), \quad i = 1, \ldots, r-1$$

where $\beta_1, \ldots, \beta_{r-1}$ are positive numbers.

Theorem 3 (Levant 1998 a, 2001): *Let system (3) have relative degree* r *with respect to the output function σ and (5) be fulfilled. Suppose also that trajectories of system (3) are infinitely extendible in time for any Lebes-*

gue-measurable bounded control. Then with properly chosen positive parameters $\beta_1, \ldots, \beta_{r-1}$, α the controller

$$u = -\alpha \operatorname{sign}(\phi_{r-1,r}(\sigma, \dot\sigma, \ldots, \sigma^{(r-1)})) \qquad (10)$$

leads to the establishment of an r-sliding mode $\sigma \equiv 0$ attracting each trajectory in finite time. The convergence time is a locally bounded function of initial conditions.

The proof of Theorem 3 is given in Appendix 1. It is the first publication of the proof. The assumption on the solution extension possibility means in practice that the system be weakly minimum phase. The positive parameters $\beta_1, \ldots, \beta_{r-1}$ are to be chosen sufficiently large in the index order. They determine a controller family applicable to all systems (3) of relative degree r satisfying (5) for some C, K_{m} and K_{M}. Parameter $\alpha > 0$ is to be chosen specifically for any fixed C, K_{m} and K_{M}. The proposed controller may be generalized in many ways. For example, coefficients of $N_{i,r}$ may be any positive numbers, equation (10) can be smoothed (Levant 1999).

Certainly, the number of choices of β_i is infinite. Here are a few examples with β_i tested for $r \le 4$, p being the least common multiple of $1, 2, \ldots, r$. The first is the relay controller, the second coincides with (7).

1. $u = -\alpha \operatorname{sign}\sigma$,

2. $u = -\alpha \operatorname{sign}(\dot\sigma + |\sigma|^{1/2}\operatorname{sign}\sigma)$

3. $u = -\alpha(\ddot\sigma + 2(|\dot\sigma|^3 + |\sigma^2|)^{1/6}\operatorname{sign}(\dot\sigma + |\sigma|^{2/3}\operatorname{sign}\sigma))$

4. $u = -\alpha\{\dddot\sigma + 3(\ddot\sigma^6 + \dot\sigma^4 + |\sigma|^3)^{1/12}\operatorname{sign}[\ddot\sigma + (\dot\sigma^4$

   $+ |\sigma|^3)^{1/6}\operatorname{sign}(\dot\sigma + 0.5|\sigma|^{3/4}\operatorname{sign}\sigma)]\}$

5. $u = -\alpha \operatorname{sign}(\sigma^{(4)} + \beta_4(|\sigma|^{12} + |\dot\sigma|^{15} + |\ddot\sigma|^{20}$

   $+ |\dddot\sigma|^{30})^{1/60}\operatorname{sign}(\dddot\sigma + \beta_3(|\sigma|^{12} + |\dot\sigma|^{15}$

   $+ |\ddot\sigma|^{20})^{1/30}\operatorname{sign}(\ddot\sigma + \beta_2(|\sigma|^{12}$

   $+ |\dot\sigma|^{15})^{1/20}\operatorname{sign}(\dot\sigma + \beta_1|\sigma|^{4/5}\operatorname{sign}\sigma))))$

Obviously, parameter α is to be taken negative with $(\partial/\partial u)\sigma^{(r)} < 0$. Controller (10) is certainly insensitive to any disturbance which keeps the relative degree and (5). No matching condition having been supposed, the residual uncertainty reveals itself in the r-sliding motion equations (in other words, in zero dynamics).

The idea of the controller is that a 1-sliding mode is established on the smooth parts of the discontinuity set Γ of (10) (figures 2, 3). That sliding mode is described by the differential equation $\phi_{r-1,r} = 0$ providing in its turn for the existence of a 1-sliding mode $\phi_{r-2,r} = 0$. But the primary sliding mode disappears at the moment when the secondary one is to appear. The resulting movement
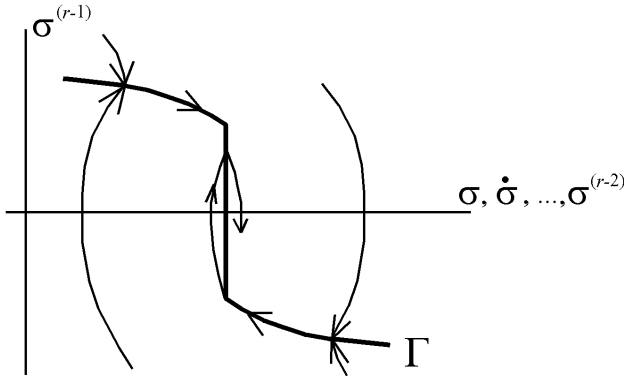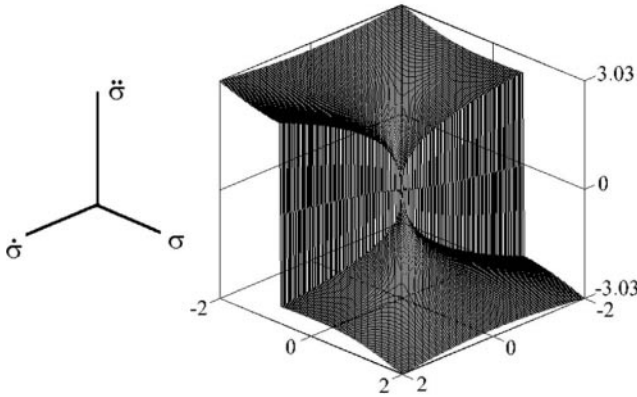
Figure 2.   The idea of the *r*-sliding controller.

Figure 3.   The 3-sliding controller discontinuity set.

takes place in some vicinity of the cylindrical subset of Γ satisfying $\phi_{r-2,r} = 0$, transfers in finite time into some vicinity of the subset satisfying $\phi_{r-3,r} = 0$ and so on. While the trajectory approaches the *r*-sliding set, set Γ retracts to the origin in the coordinates $\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}$.

Controller (10) requires the availability of $\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}$. That information demand may be lowered. Let the measurements be carried out at times t_i with constant step $\tau > 0$. Consider the controller

$$u(t) = -\alpha \operatorname{sign}\left(\Delta\sigma_i^{(r-2)} + \beta_{r-1}\tau N_{r-1,r}(\sigma_i, \dot{\sigma}_i, \ldots, \sigma_i^{(r-2)})\right.$$

$$\left. \operatorname{sign}\left(\phi_{r-2,r}(\sigma_i, \dot{\sigma}_i, \ldots, \sigma_i^{(r-2)})\right)\right) \qquad (11)$$

where $\sigma_i^{(j)} = \sigma^{(j)}(t_i, x(t_i)), \Delta\sigma_i^{(r-2)} = \sigma_i^{(r-2)} - \sigma_{i-1}^{(r-2)}$, $t \in [t_i, t_{i+1})$.

Theorem 4 (Levant 1998 b, 2001 a):   *Under conditions of Theorem 1 with discrete measurements both algorithms* (10) *and* (11) *provide in finite time for fulfilment of the inequalities*

$$|\sigma| < a_0\tau^r, |\dot{\sigma}| < a_1\tau^{r-1}, \ldots, |\sigma^{(r-1)}| < a_{r-1}\tau$$

for some positive constants $a_0, a_1, \ldots, a_{r-1}$. *The convergence time is a locally bounded function of initial conditions.*

That is the best possible accuracy attainable with discontinuous $\sigma^{(r)}$ separated from zero (Levant 1993). The proof of Theorem 4 is given in Appendix 1. Following are some remarks on the usage of the proposed controllers.

Convergence time may be reduced increasing the coefficients β_j. Another way is to substitute $\lambda^{-j}\sigma^{(j)}$ for $\sigma^{(j)}$, $\lambda^r\alpha$ for α and $\lambda\tau$ for τ in (10) and (11), $\lambda > 1$, causing convergence time to be diminished approximately by λ times. As a result the coefficients of $N_{i,r}$ will differ from 1.

Local application of the controller. In practical applications condition (5) is often invalid globally, but still holds in some restricted area of the state space containing the actual region of the system operation. In fact, that is always true, if the constraint keeping problem is well posed from the engineering point of view. The practical implementation of the controller is straight-forward in that case and is based on the following simple proposition proved in Appendix 1.

Proposition 3:   *Under the conditions of Theorem 3, for any* $R_M > 0$ *there exists such* R_m, $R_M > R_m > 0$, *that any trajectory starting in the disc of radius* R_m *which is centred at the origin of the space* $\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}$ *does not leave the larger disc of the radius* R_M *while converging to the origin. Moreover,* $R_M - R_m$ *and the convergence time can be made arbitrarily small choosing* $\beta_1, \ldots, \beta_{r-1}, \alpha$, *sufficiently large in the list order.*

Implementation of r-sliding controller when the relative degree is less than r. Let the relative degree k of the process be less than r. Introducing successive time derivatives $u, \dot{u}, \ldots, u^{(r-k-1)}$ as a new auxiliary variables and $u^{(r-k)}$ as a new control, achieve a system with relative degree r. Condition (5) is locally satisfied for the new control. The standard *r*-sliding controller may be now locally applied. The resulting control $u(t)$ is an $(r-k-1)$-smooth function of time with $k < r-1$, a Lipschitz function with $k = r-1$ and a bounded 'infinite-frequency switching' function with $k = r$. A global controller was developed by Levant (1993) for $r = 2, k = 1$.

Chattering removal. The same trick removes the chattering effect. For example, substituting $u^{(r-1)}$ for u in (10), receive a local *r*-sliding controller to be used instead of the relay controller $u = -\operatorname{sign}\sigma$ and attain the *r*th-order sliding precision with respect to τ by means of an $(r-2)$-times differentiable control with a Lipschitzian $(r-2)$th time derivative. It has to be modified like in (Levant 1993) in order to provide for the global boundedness of u and global convergence.

Controlling systems non-linear on control. Consider a system $\dot{x} = f(t, x, u)$ non-linear on control. Let

930   A. Levant

$(\partial/\partial u)\sigma^{(i)}(t, x, u) = 0$ for $i = 1, ..., r - 1$, $(\partial/\partial u)\sigma^{(r)}(t, x, u) > 0$. It is easy to check that

$$\sigma^{(r+1)} = \varLambda_u^{r+1}\sigma + \frac{\partial}{\partial u}\sigma^{(r)}\dot{u},$$

$$\varLambda_u(\cdot) = \frac{\partial}{\partial t}(\cdot) + \frac{\partial}{\partial x}(\cdot)f(t, x, u)$$

The problem is now reduced to that considered above with relative degree $r + 1$ by introducing a new auxiliary variable u and a new control $v = \dot{u}$.

Real-time control of output variables. The implementation of the above-listed r-sliding controllers requires real-time observation of the successive derivatives $\dot{\sigma}$, $\ddot{\sigma}, \ldots, \sigma^{(r-1)}$. In case system (3) is known and the full state is available, these derivatives may be directly calculated. In the real uncertainty case the derivatives are still to be real-time evaluated in some way. Thus, one would not theoretically need to know any model of the controlled process, only the relative degree and three constants from (5) were needed in order to adjust the controller. Unfortunately, the problem of successive real-time exact differentiation is usually considered as practically insoluble. Nevertheless, as is shown in the next section, the boundedness of $\sigma^{(r)}$, which follows from (4) and (5) allows robust exact estimation of $\dot{\sigma}$, $\ddot{\sigma}, \ldots, \sigma^{(r-1)}$ in real time.

5.   Arbitrary-order exact robust differentiator

Real-time differentiation is an old and well-studied problem. The main difficulty is the obvious differentiation sensitivity to input noises. The popular high-gain differentiators (Atassi and Khalil 2000) provide for an exact derivative when their gains tend to infinity. Unfortunately at the same time their sensitivity to small high-frequency noises also infinitely grows. With any finite gain values such a differentiator has also a finite bandwidth. Thus, being not exact, it is, at the same time, insensitive with respect to high-frequency noises. Such insensitivity may be considered both as advantage or disadvantage depending on the circumstances. Another drawback of the high-gain differentiators is their peaking effect: the maximal output value during the transient grows infinitely when the gains tend to infinity.

The main problem of the differentiator feedback application is the so-called separation problem. The separation principle means that a controller and an observer (differentiator) can be designed separately, so that the combined observer-controller output feedback preserve the main features of the controller with the full state available. The separation principle was proved for asymptotic continuous-feedback stabilization of autonomous systems with high-gain observers (Atassi and Khalil 2000, Isidori *et al.* 2000). That important result

is realized in spite of the non-exactness of high-gain observers with any fixed finite gain values. The qualitative explanation is that the output derivatives of all orders vanish during the continuous-feedback stabilization. Thus, the frequency of the signal to be differentiated also vanishes and the differentiator provides for asymptotically exact derivatives. On the contrary, in the case considered in the previous section $\sigma^{(r)}$ is chattering with a finite magnitude and a frequency tending to infinity while approaching the r-sliding mode. Thus, the signal σ is problematic for a high-gain differentiator. Indeed, the closer to the r-sliding mode, the higher gain is needed to produce a good derivative estimation of $\sigma^{(r-1)}$. Actually, the high-gain differentiator will differentiate only the slowly changing average output component. As a result, convergence into some vicinity of the r-sliding mode could only be attained.

The sliding-mode differentiators (Golembo *et al.* 1976, Yu and Xu 1996) also do not provide for exact differentiation with finite-time convergence due to the output filtration. The differentiator by Bartolini *et al.* (2000) is based on a 2-sliding-mode controller using the real-time measured sign of the derivative to be calculated. Therefore, the first finite difference of the differentiator input is used with the sampling step proportional to the square root of the maximal noise magnitude. That is rather inconvenient and requires possibly lacking information on the noise.

Exact derivatives may be calculated by successive implementation of a robust exact first-order differentiator (Levant 1998 a) with finite-time convergence. That differentiator is based on 2-sliding mode and is proved to feature the best possible asymptotics in the presence of infinitesimal Lebesgue-measurable measurement noises, if the second time derivative of the unknown base signal is bounded. The accuracy of that differentiator is proportional to $\varepsilon^{1/2}$, where ε is the maximal measurement-noise magnitude and is also assumed to be unknown. Therefore, having been n times successively implemented, that differentiator will provide for the nth-order differentiation accuracy of the order of $\varepsilon^{(2^{-n})}$. Thus, the differentiation accuracy deteriorates rapidly. On the other hand, it is proved by Levant (1998 a) that when the Lipschitz constant of the nth derivative of the unknown clear-of-noise signal is bounded by a given constant L, the best possible differentiation accuracy of the ith derivative is proportional to $L^{i/(n+1)}\varepsilon^{(n+1-i)/(n+1)}$, $i = 0, 1, \ldots, n$. Therefore, a special differentiator is to be designed for each differentiation order.

Let input signal $f(t)$ be a function defined on $[0, \infty)$ consisting of a bounded Lebesgue-measurable noise with unknown features and an unknown base signal $f_0(t)$ with the nth derivative having a known Lipschitz constant $L > 0$. The problem is to find real-time robust

estimations of $\dot{f}_0(t), \ddot{f}_0(t), \ldots, f_0^{(n)}(t)$ being exact in the absence of measurement noises.

Two similar recursive schemes of the differentiator are proposed here. Let an $(n-1)$th-order differentiator $D_{n-1}(f(\cdot), L)$ produce outputs $D_{n-1}^i, i = 0, 1, \ldots, n-1$, being estimations of $f_0, \dot{f}_0, \ddot{f}_0, \ldots, f_0^{(n-1)}$ for any input $f(t)$ with $f_0^{(n-1)}$ having Lipschitz constant $L > 0$. Then the nth-order differentiator has the outputs $z_i = D_n^i$, $i = 0, 1, \ldots, n$, defined as

$$\left. \begin{aligned} &\dot{z}_0 = v, \quad v = -\lambda_0 |z_0 - f(t)|^{n/(n+1)} \text{sign}(z_0 - f(t)) + z_1 \\ &z_1 = D_{n-1}^0(v(\cdot), L), \ldots, z_n = D_{n-1}^{n-1}(v(\cdot), L) \end{aligned} \right\}$$
(12)

Here $D_0(f(\cdot), L)$ is a simple non-linear filter

$$D_0 : \dot{z} = -\lambda \, \text{sign}(z - f(t)), \qquad \lambda > L \qquad (13)$$

Thus, the first-order differentiator coincides here with the above-mentioned differentiator (Levant 1998 a) :

$$\left. \begin{aligned} &\dot{z}_0 = v, \quad v = -\lambda_0 |z_0 - f(t)|^{1/2} \text{sign}(z_0 - f(t)) + z_1 \\ &\dot{z}_1 = -\lambda_1 \text{sign}(z_1 - v) = -\lambda_1 \text{sign}(z_0 - f(t)) \end{aligned} \right\}$$
(14)

Another recursive scheme is based on the differentiator (14) as the basic one. Let $\tilde{D}_{n-1}(f(\cdot), L)$ be such a new $(n-1)$th-order differentiator, $n \geq 1$, $\tilde{D}_1(f(\cdot), L)$ coinciding with the differentiator $D_1(f(\cdot), L))$ given by (14). Then the new scheme is defined as

$$\left. \begin{aligned} &\dot{z}_0 = v \\ &v = -\lambda_0 |z_0 - f(t)|^{n/(n+1)} \text{sign}(z_0 - f(t)) + w_0 + z_1 \\ &\dot{w}_0 = -\alpha_0 |z_0 - f(t)|^{(n-1)/(n+1)} \text{sign}(z_0 - f(t)) \\ &z_1 = \tilde{D}_{n-1}^0(v(\cdot), L), \ldots, z_n = \tilde{D}_{n-1}^{n-1}(v(\cdot), L) \end{aligned} \right\}$$
(15)

The resulting 2nd-order differentiator (Levant 1999) is

$$\dot{z}_0 = v_0, \quad v_0 = -\lambda_0 |z_0 - f(t)|^{2/3} \text{sign}(z_0 - f(t)) + w_0 + z_1$$

$$\dot{w}_0 = -\alpha_0 |z_0 - f(t)|^{1/3} \text{sign}(z_0 - f(t))$$

$$\dot{z}_1 = v_1, \quad v_1 = -\lambda_1 |z_1 - v_0|^{1/2} \text{sign}(z_1 - v_0) + w_1$$

$$\dot{w}_1 = -\alpha_1 \text{sign}(z_1 - v_0); \quad z_2 = w_1$$

Similarly, a 2nd-order differentiator from each of these sequences may be used as a base for a new recursive scheme. An infinite number of differentiator schemes may be constructed in this way. The only requirement is that the resulting systems be homogeneous in a sense described further. While the author

has checked only the above two schemes (12), (13) and (14), (15), the conjecture is that all such schemes produce working differentiators, provided suitable parameter choice. Differentiator (12) takes on the form

$$\left. \begin{aligned} &\dot{z}_0 = v_0, \\ &v_0 = -\lambda_0 |z_0 - f(t)|^{n/(n+1)} \text{sign}(z_0 - f(t)) + z_1 \\ &\dot{z}_1 = v_1, \quad v_1 = -\lambda_1 |z_1 - v_0|^{(n-1)/n} \text{sign}(z_1 - v_0) + z_2 \\ &\vdots \\ &\dot{z}_{n-1} = v_{n-1}, \\ &v_{n-1} = -\lambda_{n-1} |z_{n-1} - v_{n-2}|^{1/2} \text{sign}(z_{n-1} - v_{n-2}) + z_n \\ &\dot{z}_n = -\lambda_n \text{sign}(z_n - v_{n-1}) \end{aligned} \right\}$$
(16)

Theorem 5: *The parameters being properly chosen, the following equalities are true in the absence of input noises after a finite time of a transient process*

$$z_0 = f_0(t); \quad z_i = v_{i-1} = f_0^{(i)}(t), \quad i = 1, \ldots, n$$

Moreover, the corresponding solutions of the dynamic systems are Lyapunov stable, i.e. finite-time stable (Rosier 1992). The theorem means that the equalities $z_i = f_0^{(i)}(t)$ are kept in 2-sliding mode, $i = 0, \ldots, n-1$. Here and further all Theorems are proved in Appendix 2.

Theorem 6: *Let the input noise satisfy the inequality $|f(t) - f_0(t)| \leq \varepsilon$. Then the following inequalities are established in finite time for some positive constants μ_i, ν_i depending exclusively on the parameters of the differentiator*

$$|z_i - f_0^{(i)}(t)| \leq \mu_i \varepsilon^{(n-i+1)/(n+1)}, \quad i = 0, \ldots, n$$

$$|vi - f_0^{(i+1)}(t)| \leq \nu_i \varepsilon^{(n-i)/(n+1)}, \quad i = 0, \ldots, n-1$$

Consider the discrete-sampling case, when $z_0(t_j) - f(t_j)$ is substituted for $z_0 - f(t)$ with $t_j \leq t < t_{j+1}, t_{j+1} - t_j = \tau > 0$.

Theorem 7: *Let $\tau > 0$ be the constant input sampling interval in the absence of noises. Then the following inequalities are established in finite time for some positive constants μ_i, ν_i depending exclusively on the parameters of the differentiator*

$$|z_i - f_0^{(i)}(t)| \le \mu_i \tau^{n-i+1}, \quad i = 0,\dots,n$$

$$|v_i - f_0^{(i+1)}(t)| \le \nu_i \tau^{n-i}, \quad i = 0,\dots,n-1$$

In particular, the nth derivative error is proportional to τ. The latter theorem means that there are a number of real sliding modes of different orders. Nevertheless, nothing can be said on the derivatives of v_i and of $\dot{z}_i$, because they are not continuously differentiable functions.

Homogeneity of the differentiators. The differentiators are invariant with respect to the transformation

$$(t,f,z_i,v_i,w_i) \mapsto (\eta t, \eta^{n+1}f, \eta^{n-i+1}z_i, \eta^{n-i}v_i, \eta^{n-i}w_i)$$

The parameters α_i, λ_i are to be chosen recursively in such a way that $\alpha_1, \lambda_1, \dots, \alpha_n, \lambda_n, \lambda_n$ provide for the convergence of the $(n-1)$th-order differentiator with the same Lipschitz constant L, and α_0, λ_0 be sufficiently large (α_0 is chosen first). The best way is to choose them by computer simulation. A choice of the 5th-order differentiator parameters with $L = 1$ is demonstrated in §7. Recall that it contains parameters of all lower-order differentiators. Substituting $f(t)/L$ for $f(t)$ and taking new coordinates $z_i' = Lz_i$, $v_i' = Lv_i$, $w_i' = Lw_i$ achieve the following proposition.

Proposition 4: *Let parameters α_{0i}, $\lambda_{0i}, i = 0,1,\dots,n$, of differentiators (12), (13) or (14), (15) provide for exact nth-order differentiation with $L = 1$. Then the parameters $\alpha_i = \alpha_{0i}L^{2/(n-i+1)}$, $\lambda_i = \lambda_{0i}L^{1/(n-i+1)}$ are valid for any $L > 0$ and provide for the accuracy $|z_i - f_0^{(i)}(t)| \le \mu_i L^{i/(n+1)}\varepsilon^{(n-i+1)/(n+1)}$ for some $\mu_i \ge 1$.*

The separation principle is trivially fulfilled for the proposed differentiator. Indeed, the differentiator being exact, the only requirements for its implementation are the boundedness of some higher-order derivative of its input and the impossibility of the finite-time escape during the differentiator transient. Hence, the differentiator may be used in almost any feedback. Mark that the differentiator transient may be made arbitrarily short by means of the parameter transformation from Proposition 4, and the differentiator does not feature peaking effect (see the proof of Theorem 5 in the Appendices).

Remarks: It is easy to see that the kth-order differentiator provides for a much better accuracy of the lth derivative, $l < k$, than the lth-order differentiator (Theorem 6). A similar idea is realized to improve the first derivative by Krupp *et al.* (2001). It is easy to check that after exclusion of the variables v_i differentiator (16) may be rewritten in the non-recursive form

$$\dot{z}_0 = -\kappa_0|z_0 - f(t)|^{n/(n+1)}\text{sign}\,(z_0 - f(t)) + z_1$$

$$\dot{z}_i = -\kappa_i|z_0 - f(t)|^{(n-i)/(n+1)}\text{sign}\,(z_0 - f(t)) + z_{i+1},$$

$$i = 1,\dots,n-1$$

$$\dot{z}_n = -\kappa_n \text{sign}\,(z_0 - f(t))$$

for some positive κ_i calculated on the basis of $\lambda_0 \dots, \lambda_n$.

6. Universal output-feedback SISO controller

The results of this section have been just recently obtained (Levant 2002), and the author supposes to describe them in a special paper. Thus, only a brief description is provided. Consider uncertain system (3), (5). Combining controller (10) and differentiator (16) achieve a combined single-input–single-output (SISO) controller

$$u = -\alpha \text{sign}\,(\phi_{r-1,r}(z_0,z_1,\dots,z_{r-1}))$$

$$\dot{z}_0 = v_0, \quad v_0 = -\lambda_{0,0}L^{1/r}|z_0 - \sigma|^{(r-1)/r}\text{sign}\,(z_0 - \sigma) + z_1$$

$$\dot{z}_1 = v_1,$$

$$v_1 = -\lambda_{0,1}L^{1/(r-1)}|z_1 - v_0|^{(r-2)/(r-1)}\text{sign}\,(z_1 - v_0) + z_2$$

$$\vdots$$

$$\dot{z}_{r-2} = v_{r-2},$$

$$v_{r-2} = -\lambda_{0,r-2}L^{1/2}|z_{r-2} - v_{r-3}|^{1/2}\text{sign}\,(z_{r-2} - v_{r-3}) + z_{r-1}$$

$$\dot{z}_{r-1} = -\lambda_{0,r-1}L\,\text{sign}\,(z_{r-1} - v_{r-2})$$

where parameters $\lambda_i = \lambda_{0,i}L^{1/(r-I)}$ of the differentiator are chosen according to the condition $|\sigma^{(r)}| \le L, L \ge C + \alpha K_\text{M}$. In their turn, parameters λ_{0i} are chosen in advance for $L = 1$ (Proposition 4). Thus, parameters of controller (10) are chosen separately of the differentiator. In case when C and K_M are known, only one parameter α is really needed to be tuned, otherwise both L and α might be found in computer simulation. Theorems 3 and 4 hold also for the combined output-feedback controller. In particular, under the conditions of Theorem 3 the combined controller provides for the global convergence to the r-sliding mode $\sigma \equiv 0$ with the transient time being a locally bounded function of the initial conditions.

On the other hand, let the initial conditions of the differentiator belong to some compact set. Then for any two embedded discs centred at the origin of the space $\sigma, \dot{\sigma}, \dots, \sigma^{(r-1)}$ the parameters of the combined controller can be chosen in such a way that all trajectories starting in the smaller disc do not leave the larger disc during their finite-time convergence to the origin. The

convergence time can be made arbitrarily small. That allows for the local controller application.

With discrete measurements, in the absence of input noises, the controller provides for the rth-order real sliding $\sup|\sigma| \sim \tau^r$, where τ is the sampling interval. Therefore, the differentiator does not spoil the r-sliding asymptotics if the input noises are absent. It is also proved that the resulting controller is robust and provides for the accuracy proportional to the maximal error of the input measurement (the input noise magnitude). Note once more that the proposed controller does not require detailed mathematical model of the process to be known.

7. Simulation examples

7.1. *Numeric differentiation*

Following are equations of the 5th-order differentiator with simulation-tested coefficients for $L = 1$

$$\dot{z}_0 = v_0, \ v_0 = -12|z_0 - f(t)|^{5/6}\mathrm{sign}\,(z_0 - f(t)) + z_1 \quad (17)$$

$$\dot{z}_1 = v_1, \ v_1 = -8|z_1 - v_0|^{4/5}\mathrm{sign}\,(z_1 - v_0) + z_2 \quad (18)$$

$$\dot{z}_2 = v_2, \ v_2 = -5|z_2 - v_1|^{3/4}\mathrm{sign}\,(z_2 - v_1) + z_3 \quad (19)$$

$$\dot{z}_3 = v_3, \ v_3 = -3|z_3 - v_2|^{2/3}\mathrm{sign}\,(z_3 - v_2) + z_4 \quad (20)$$

$$\dot{z}_4 = v_4, \ v_4 = -1.5|z_4 - v_3|^{1/2}\mathrm{sign}\,(z_4 - v_3) + z_5 \quad (21)$$

$$\dot{z}_5 = -1.1\,\mathrm{sign}\,(z_5 - v_4) \quad (22)$$

The differentiator parameters can be easily changed, for it is not very sensitive to their values. The tradeoff is as follows: the larger the parameters, the faster the convergence and the higher sensitivity to input noises and the sampling step.

As mentioned, differentiator (17)–(22) also contains differentiators of the lower orders. For example, according to Proposition 4, the second-order differentiator for the input f with $|\dddot{f}| \le L$ takes on the form

$$\dot{z}_0 = v_0, \quad v_0 = -3L^{1/3}|z_0 - f|^{2/3}\,\mathrm{sign}\,(z_0 - f) + z_1$$

$$\dot{z}_1 = v_1, \quad v_1 = -1.5L^{1/2}|z_1 - v_0|^{1/2}\,\mathrm{sign}\,(z_1 - v_0) + z_2$$

$$\dot{z}_2 = -1.1L\,\mathrm{sign}\,(z_2 - v_1)$$

Differentiator (17)–(22) and its 3rd-order sub-differentiator (19)–(22) (differentiating here the internal variable v_1) were used for simulation. Initial values of the differentiator state were taken zero with exception for the initial estimation z_0 of f, which is taken equal to the initial measured value of f. The base input signal

$$f_0(t) = 0.5\sin 0.5t + 0.5\cos t \quad (23)$$

was taken for the differentiator testing. Derivatives of $f_0(t)$ do not exceed 1 in absolute value.

Third-order differentiator. The measurement step $\tau = 10^{-3}$ was taken, noises are absent. The attained accuracies are 5.8×10^{-12}, 1.4×10^{-8}, 1.0×10^{-5} and 0.0031 for the signal tracking, the first, second and third derivatives respectively. The derivative tracking deviations changed to 8.3×10^{-16}, 1.8×10^{-11}, 1.2×10^{-7} and 0.00036 respectively after τ was reduced to 10^{-4}. That corresponds to Theorem 7.

Fifth-order differentiator. The attained accuracies are 1.1×10^{-16}, 1.29×10^{-12}, 7.87×10^{-10}, 5.3×10^{-7}, 2.0×10^{-4} and 0.014 for tracking the signal, the first, second, third, fourth and fifth derivatives respectively with $\tau = 10^{-4}$ (figure 4(*a*)). *There is no significant improvement with further reduction of τ.* The author wanted to demonstrate the 10th-order differentiation, but found that differentiation of the order exceeding 5 is unlikely to be performed with the standard software. Further calculations are to be carried out with precision higher than the standard *long double* precision (128 bits per number).
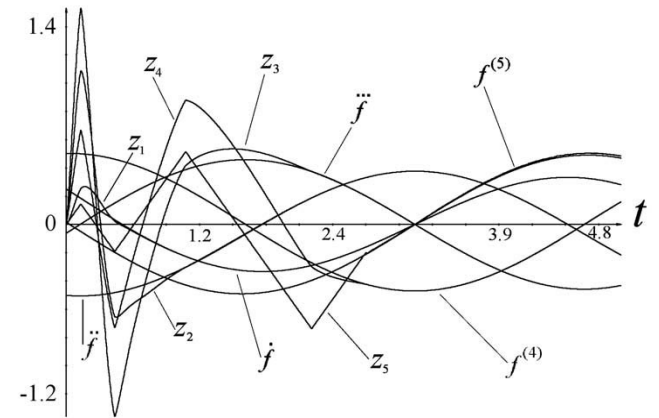
a. Signal without noise
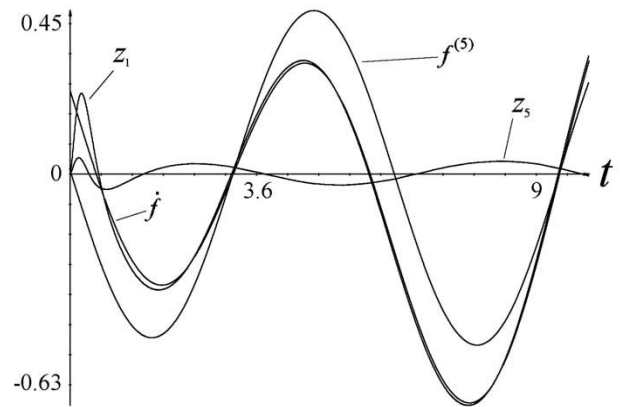
b. Noisy signal

Figure 4. Fifth-order differentiation.

Sensitivity to noises. The main problem of the differentiation is certainly its well-known sensitivity to noises. As we have seen, even small computer calculation errors appear to be a considerable noise in the calculation of the fifth derivative. Recall that, when the nth derivative has the Lipschitz constant 1 and the noise magnitude is ε, the best possible accuracy of the ith-order differentiation, $i \leq n$, is $k(i,n)\varepsilon^{(n-i+1)/(n+1)}$ (Levant 1998 a), where $k(i,n) > 1$ is a constant independent on the differentiation realization. That is a minimax (worst case) evaluation. Since differentiator (17)–(22) assumes this Lipschitz input condition, it satisfies this accuracy restriction as well (see also Theorems 6 and 7). In particular, with the noise magnitude $\varepsilon = 10^{-6}$ the maximal 5th derivative error exceeds $\varepsilon^{1/6} = 0.1$. For comparison, if the successive first-order differentiation were used, the respective maximal error would be at least $\varepsilon^{(2-5)} = 0.649$ and some additional conditions on the input signal would be required. Taking 10% as a border, achieve that the direct successive differentiation does not give reliable results starting with the order 3, while the proposed differentiator may be used up to the order 5.

With the noise magnitude 0.01 and the noise frequency about 1000 the 5th-order differentiator produces estimation errors 0.000 42, 0.0088, 0.076, 0.20, 0.34 and 0.52 for signal (23) and its five derivatives respectively (figure 4(*b*)). The differentiator performance does not significantly depend on the noise frequency. The author found that the second differentiation scheme (15) provides for slightly better accuracies.

7.2. *Output-feedback control simulation*

Consider a simple kinematic model of car control (Murray and Sastry 1993)

$$\dot{x} = v \cos \varphi, \qquad \dot{y} = v \sin \varphi$$

$$\dot{\varphi} = (v/l) \tan \theta$$

$$\dot{\theta} = u$$

where x and y are Cartesian coordinates of the rear-axle middle point, φ is the orientation angle, v is the longitudinal velocity, l is the length between the two axles and θ is the steering angle (figure 5). The task is to steer the car from a given initial position to the trajectory $y = g(x)$, while $x, g(x)$ and y are assumed to be measured in real time. Note that the actual control here is θ and $\dot{\theta} = u$ is used as a new control in order to avoid discontinuities of θ. Any practical implementation of the developed here controller would require some real-time coordinate transformation with θ approaching $\pm \pi/2$. Define
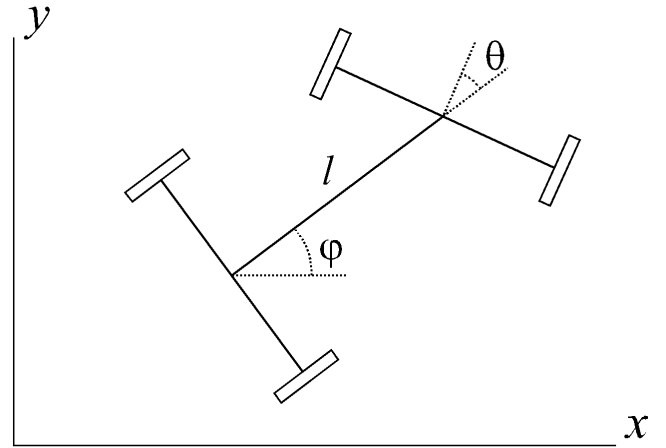
$$\sigma = y - g(x) \qquad (24)$$

Figure 5.   Kinematic car model.

Let $v = \text{const} = 10\,\text{m/s}$, $l = 5\,\text{m}$, $g(x) = 10 \sin(0.05x) + 5$, $x = y = \varphi = \theta = 0$ at $t = 0$. The relative degree of the system is 3 and the listed 3-sliding controller may be applied here. The resulting steering angle dependence on time is not sufficiently smooth (Levant 2001 b), therefore the relative degree is artificially increased up to 4, $\dot{u}$ having been considered as a new control. The 4-sliding controller from the list in §3 is applied now, $\alpha = 20$ is taken. The following 3rd order differentiator was implemented:

$$\dot{z}_0 = v_0, \quad v_0 = -25|z_0 - \sigma|^{3/4} \operatorname{sign}(z_0 - \sigma) + z_1 \qquad (25)$$

$$\dot{z}_1 = v_1, \quad v_1 = -25|z_1 - v_0|^{2/3} \operatorname{sign}(z_1 - v_0) + z_2 \qquad (26)$$

$$\dot{z}_2 = v_2, \quad v_2 = -33|z_2 - v_1|^{1/2} \operatorname{sign}(z_2 - v_1) + z_3 \qquad (27)$$

$$\dot{z}_3 = -500 \operatorname{sign}(z_3 - v_2) \qquad (28)$$

The coefficient in (28) is large due to the large values of $\sigma^{(4)}$, other coefficients were taken according to Proposition 4 and (17)–(22). During the first half-second the control is not applied in order to allow the convergence of the differentiator. Substituting z_0, z_1, z_2 and z_3 for σ, $\dot{\sigma}$, $\ddot{\sigma}$ and $\dddot{\sigma}$ respectively, obtain the following 4-sliding controller

$$u = 0, \quad 0 \leq t < 0.5$$

$$u = -20 \operatorname{sign} \{z_3 + 3(z_2^6 + z_1^4 + |z_0|^3)^{1/12} \operatorname{sign}[z_2$$
$$+ (z_1^4 + |z_0|^3)^{1/6} \operatorname{sign}(z_1 + 0.5|z_0|^{3/4} \operatorname{sign} z_0)]\}$$

$$t \leq 0.5$$

The trajectory and function $y = g(x)$ with the sampling step $\tau = 10^{-4}$ are shown in figure 6(*a*). The integration was carried out according to the Euler method, the only method reliable with sliding-mode simulation. Graphs of σ, $\dot{\sigma}$, $\ddot{\sigma}$, $\dddot{\sigma}$ are shown in figure

$6(b)$. The differentiator performance within the first $1.5\,\text{s}$ is demonstrated in figure $6(c)$. The steering angle graph (actual control) is presented in figure $6(d)$. The sliding accuracies $|\sigma| \le 9.3 \times 10^{-8}$, $|\dot{\sigma}| \le 7.8 \times 10^{-5}$, $|\ddot{\sigma}| \le 6.6 \times 10^{-4}$, $|\dddot{\sigma}| \le 0.43$ were attained with the sampling time $\tau = 10^{-4}$.

8. Conclusions and discussion of the obtained results

Arbitrary-order real-time exact differentiation together with the arbitrary-order sliding controllers provide for full SISO control based on the input measurements only, when the only information on the controlled uncertain weakly-minimum-phase process is actually its relative degree.

A family of r-sliding controllers with finite time convergence is presented for any natural number r, providing for the full real-time control of the output variable if the relative degree r of the dynamic system is constant and known. Whereas 1- and 2-sliding modes were used mainly to keep auxiliary constraints, arbitrary-order sliding controllers may be considered as general-purpose controllers. In case the mathematical model of the system is known and the full state is available, the real-time derivatives of the output variable are directly calculated, and the controller implementation is straightforward and does not require reduction of the dynamic system to any specific form. If boundedness restrictions (5) are globally satisfied, the control is also global and the input is globally bounded. Otherwise the controller is still locally applicable.

In the uncertainty case a detailed mathematical model of the process is not needed. Necessary time derivatives of the output can be obtained by means of the proposed robust exact differentiator with finite-time convergence. The proposed differentiator allows real-time robust exact differentiation up to any given order l, provided the next $(l+1)$th derivative is bounded by a known constant. These features allow wide application of the differentiator in non-linear control theory. Indeed, after finite time transient in the absence of input noises its outputs can be considered as *exact direct measurements* of the derivatives. Therefore, the separation principle is trivially true for almost any feedback. At the same time, in the presence of measurement noises the differentiation accuracy inevitably deteriorates rapidly with the growth of the differentiation order (Levant 1998 a), and direct observation of the derivatives is preferable. The exact derivative estimation does not require tending some parameters to infinity or to zero. Even when treating noisy signals, the differentiator

a. Car trajectory

b. 4-sliding deviations

c. Differentiator convergence
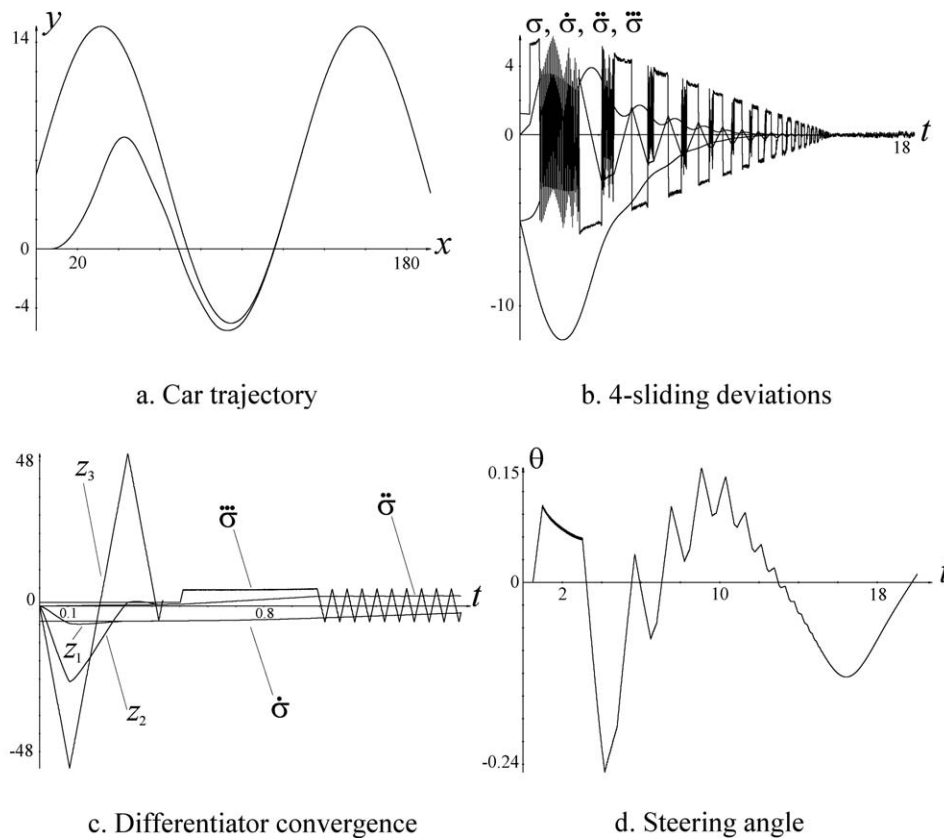
d. Steering angle

Figure 6.   4-sliding car control.

performance only improves with the sampling step reduction.

The resulting controller provides for extremely high tracking accuracy in the absence of noises. The sliding accuracy is proportional to τ^r, τ being a sampling period and r being the relative degree. That is the best possible accuracy with discontinuous rth derivative of the output (Levant 1993). It may be further improved increasing the relative degree artificially, which produces arbitrarily smooth control and removes also the chattering effect.

The proposed controllers are easily developed for any relative degree, at the same time most of the practically important problems in output control are covered by the cases when relative degree equals 2, 3, 4 and 5. Indeed, according to the Newton law, the relative degree of a spatial variable with respect to a force, being understood as a control, is 2. Taking into account some dynamic actuators, achieve relative degree 3 or 4. If the actuator input is required to be a continuous Lipschitz function, the relative degree is artificially increased to 4 or 5. Recent results (Bartolini *et al.* 1999 b) seem to allow the implementation of the developed controllers for general multi-input–multi-output systems.

Appendix 1.   Proofs of Theorems 3, 4 and Proposition 3

The general idea of the proofs is presented in §4 and is illustrated by figures 2 and 3.

Preliminary notions: The following notions are needed to understand the proof. They are based on results by Filippov (1988).

Differential inclusion $\dot{\xi} \in \Xi(\xi)$, $\xi \in \mathbf{R}^m$ is called further Filippov inclusion if for any ξ:

1. $\Xi(\xi)$ is a closed non-empty convex set;

2. $\Xi(\xi) \subset \{v \in \mathbf{R}^m | \|v\| \leq \rho(\xi)\}$, where $\rho(\xi)$ is a continuous function;

3. the maximal distance of the points of $\Xi(\xi')$ from $\Xi(\xi)$ tends to zero when $\xi' \to \xi$.

Recall that any solution of a differential inclusion is an absolutely continuous function satisfying the inclusion almost everywhere, and that any differential equation with a discontinuous right-hand side is understood as equivalent to some Filippov inclusion.

The graph of a differential inclusion $\dot{\xi} \in \Xi(\xi)$, $\xi \in \mathbf{R}^m$ is the set $\{(\xi, \dot{\xi}) \in \mathbf{R}^m \times \mathbf{R}^m | \dot{\xi} \in \Xi(\xi)\}$. A differential inclusion $\dot{\xi} \in \Xi'(\xi)$ is called ε-close to the Filippov inclusion $\dot{\xi} \in \Xi(\xi)$ in some region if any point of the graph of $\dot{\xi} \in \Xi'(\xi)$ is distanced by not more than ε from the graph of $\dot{\xi} \in \Xi(\xi)$. It is known that within any compact region solutions of $\dot{\xi} \in \Xi'(\xi)$ tend to some solutions of $\dot{\xi} \in \Xi(\xi)$ uniformly on any finite time interval with $\varepsilon \to 0$. In the

special case $\xi^{(m)} \in \Xi(\xi, \dot{\xi}, \ldots, \xi^{(m-1)})$, $\xi \in \mathbf{R}$, which is considered in the present paper, the graph may be considered as a set from $\mathbf{R}^m \times \mathbf{R}$. An inclusion $\xi^{(m)} \in \Xi'(\xi, \dot{\xi}, \ldots, \xi^{(m-1)})$ corresponding to the closed ε-vicinity of that graph is further called the ε-swollen inclusion. It is easy to see that this is a Filippov inclusion. An ε-swollen differential equation is the inclusion corresponding to the ε-vicinity of the corresponding Filippov inclusion.

Proof of Theorem 3:   Consider the motion of a projection trajectory of (3) and (10) in coordinates σ, $\dot{\sigma}, \ldots, \sigma^{(r-1)}$

$$\sigma^{(r)} = L_{a_e}^r \sigma(t, x) + u \frac{\partial}{\partial u} \sigma^{(r)}(t, x, u) \qquad (29)$$

Taking into account (5) achieve a differential inclusion

$$\sigma^{(r)} \in [-C, C] + [K_m, K_M] u \qquad (30)$$

which will be considered from now on instead of the real equality (29). The operations on sets are naturally understood here as sets of operation results for all possible combinations of the operand set elements. Control u is given by (10) or (11).

A given point P is called here a discontinuity point of a given function, if for any point set N of zero measure and any vicinity O of P there are at least two different limit values of the function when point $p \in O/N$ approaches P. Let Γ be the closure of the discontinuity set of sign $(\phi_{r-1,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}))$ or, in other words, of control (10).

Lemma 1:   *Set Γ partitions the whole space $\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}$ into two connected open components satisfying $\phi_{r-1,r}(\sigma, \ldots, \sigma^{(r-1)}) > 0$ and $\phi_{r-1,r}(\sigma, \ldots, \sigma^{(r-1)}) < 0$ respectively. Any curve connecting points from different components has a non-empty intersection with Γ.*

Proof: Consider any equation $\phi_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i)}) = 0$, $i = 1, \ldots, r-1$. It may be rewritten in the form

$$\sigma^{(i)} = \xi_i(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)})$$

$$= -\beta_i N_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \operatorname{sign}(\phi_{i-1,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}))$$

Let Σ_i be the closure of the discontinuity set of sign $\phi_{i,r}$, $i = 0, \ldots, r-1$, $\Gamma = \Sigma_{r-1}$, and $\Sigma_0 = \{0\} \subset \mathbf{R}$ (figure 7). Each set Σ_i lies in the space $\sigma, \dot{\sigma}, \ldots, \sigma^{(i)}$ and is, actually, a modification of the graph of $\sigma^{(i)} = \xi_i(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)})$.

The Lemma is proved by the induction principle. Obviously, Σ_0 partitions $\mathbf{R}$ into two open connected components. Let Σ_{i-1} divide the space $\mathbf{R}^i$ with coordinates $\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}$ into two open connected compon-

ents Ω_{i-1}^{+} and Ω_{i-1}^{-} with $\phi_{i-1,r} > 0$ and $\phi_{i-1,r} < 0$ respectively. It is easy to see (figure 7) that

$$\Sigma_i = \{\sigma, \dot{\sigma}, \ldots, \sigma^{(i)} | \, |\sigma^{(i)}|$$

$$\leq \beta_i N_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \,\&\, (\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \in \Sigma_{i-1} \vee$$

$$(\sigma^{(i)} = -\beta_i N_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \operatorname{sign}(\phi_{i-1,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)})))$$

$$\&\, (\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \notin \Sigma_{i-1})\}$$

$$\Omega_i^{+} = \{\sigma, \dot{\sigma}, \ldots, \sigma^{(i)} | \sigma^{(i)} > \beta_i N_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \vee$$

$$(\sigma^{(i)} > -\beta_i N_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)})$$

$$\&\, (\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \in \Omega_{i-1}^{+})\}$$

$$\Omega_i^{-} = \{\sigma, \dot{\sigma}, \ldots, \sigma^{(i)} | \sigma^{(i)} < -\beta_i N_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \vee$$

$$(\sigma^{(i)} < \beta_i N_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)})$$

$$\&\, (\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \in \Omega_{i-1}^{-})\}$$

Considering the division of the whole space $\sigma, \dot{\sigma}, \ldots, \sigma^{(i)}$ in the sets $|\sigma^{(i)}| \leq \beta_i N_{i,r}$, $\sigma^{(i)} < -\beta_i N_{i,r}$ and $\sigma^{(i)} > \beta_i N_{i,r}$ it is easily proved that Ω_i^{+} and Ω_i^{-} are connected and open. $\qquad\square$

Consider the transformation $G_\nu \colon (t, \sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}) \mapsto (\nu t, \nu^r \sigma, \nu^{r-1} \dot{\sigma}, \ldots, \nu \sigma^{(r-1)})$. It is easy to see that this linear transformation complies with understanding $\sigma^{(j)}$ as
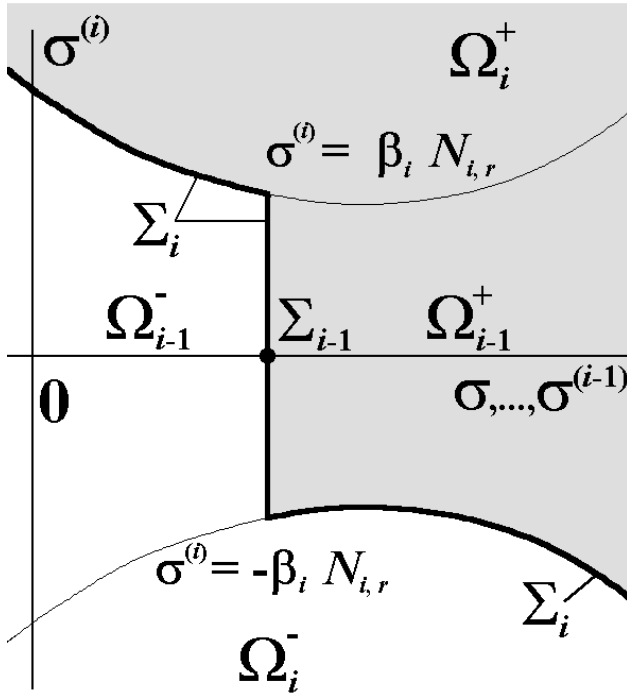

Figure 7. Internal structure of the discontinuity set.

coordinates and as derivatives as well. It is also easy to check that inclusion (30) and (10) is invariant with respect to $G_\nu, \nu > 0$. That invariance implies that any statement invariant with respect to G_ν is globally true if it is true on some set E satisfying the condition $\cup_{\nu \geq 0} G_\nu E = \mathbf{R}^r$. That reasoning is called further 'homogeneity reasoning'. For example, it is sufficient to prove the following lemma only for trajectories with initial conditions close to the origin.

Lemma 2: *With sufficiently large α any trajectory of inclusion* (30) *and* (10) *hits Γ in finite time.*

Proof: Obviously, it takes finite time to reach the region $|\sigma^{(r-1)}| \leq \beta_{r-1} N_{r-1,r}$. It is also easy to check that it takes finite time for any trajectory in a sufficiently small vicinity of the origin to cross the entire region $|\sigma^{(r-1)}| \leq \beta_{r-1} N_{r-1,r}$ if no switching happens. Thus, according to Lemma 1 set Γ is also encountered on the way. The homogeneity reasoning completes the proof. $\qquad\square$

Lemma 3: *There is a 1-sliding mode on $\phi_{r-1,r}$ in the continuity points of $\phi_{r-1,r}$ with sufficiently large α. There is such a choice of β_j that each differential equation $\phi_{i,r} = 0, i = 1, \ldots, r-1$ provides for the existence of a 1-sliding mode in the space $\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}$ on the manifold $\sigma^{(i-1)} = \xi_{i-1}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-2)})$ in continuity points of ξ_{i-1}.*

Proof: It is needed to prove that $\dot{\varphi}_{r-1,r} \operatorname{sign} \phi_{r-1,r} < \text{const} < 0$ with $\phi_{r-1,r} = 0$ on trajectories of (30) and (10). The second statement means that for each $i = 1, \ldots, r-1$ the inequalities $\dot{\varphi}_{i-1,r} \operatorname{sign} \phi_{i-1,r} < \text{const} \cdot N_{i,r} < 0$ hold with $\phi_{i-1,r} = 0$ and $\phi_{i,r} = 0$, where the latter equation is understood as a differential one and the previous determines a manifold in its state space. The proofs are quite similar. Here is an outline of the first-statement proof.

Actually it is needed to prove that $\dot{N}_{r-1,r}$ is bounded with $\phi_{r-1,r} = 0$.

$$\dot{N}_{r-1,r} = \frac{\mathrm{d}}{\mathrm{d}t} (|\sigma|^{p/r} + |\dot{\sigma}|^{p/(r-1)} + \cdots + |\sigma^{(r-2)}|^{p/2})^{1/p}$$

$$= \frac{1}{p} \left[\frac{\mathrm{d}}{\mathrm{d}t} (|\sigma|^{p/r} + |\dot{\sigma}|^{p/(r-1)} + \cdots + |\sigma^{(r-2)}|^{p/2}) \right] \Big/$$

$$(|\sigma|^{p/r} + |\dot{\sigma}|^{p/(r-1)} + \cdots + |\sigma^{(r-2)}|^{p/2})^{(p-1)/p}$$

Consider one term. The following estimation requires $p \geq r$, also the trivial inequality $(a+b)^{\mu+\nu} \geq a^\mu b^\nu, a, \mu, b, \nu > 0$, is used. Let $j = 0, 1, \ldots, r-3$, then

$$\left| \frac{\mathrm{d}}{\mathrm{d}t} |\sigma^{(j)}|^{p/(r-j)} / (|\sigma|^{p/r} + |\dot{\sigma}|^{p/(r-1)} + \cdots + |\sigma^{(r-2)}|^{p/2})^{(p-1)/p} \right|$$

$$= p/(r-j)|\sigma^{(j)}|^{p/(r-j)-1}|\sigma^{(j+1)}| \operatorname{sign} \sigma| / (|\sigma|^{p/r} + |\dot{\sigma}|^{p/(r-1)}$$

$$+ \cdots + |\sigma^{(r-2)}|^{p/2})^{(p-1)/p}$$

$$\leq p/(r-j)|\sigma^{(j)}|^{p/(r-j)-1} / |\sigma^{(j)}|^{[p/(r-j)][(p-1-(r-j-1))/p]}|\sigma^{(j+1)}| /$$

$$|\sigma^{(j+1)}|^{[p/(r-j-1)][(r-j-1)/p]}$$

$$= p/(r-j)|\sigma^{(j)}|^{(p-r+j))/(r-j)} / |\sigma^{(j)}|^{(p-r+j))/(r-j)}|\sigma^{(j+1)}| / |\sigma^{(j+1)}|$$

$$= p/(r-j)$$

With $j = r-2$ the equality $|\sigma^{(r-1)}| = \beta_{r-1}(|\sigma|^{p/r} + |\dot{\sigma}|^{p/(r-1)} + \cdots + |\sigma^{(r-2)}|^{p/2})^{1/p}$ is to be used. $\square$

Let Γ_i be the corresponding subset of Γ which is projected into $\Sigma_i, \Gamma = \Gamma_{r-1} = \Sigma_{r-1}$. Denote

$$\gamma_i = \{\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)} | |\sigma^{(j)}| \leq \beta_i N_{j,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(j-1)}) j$$

$$= i+1, i+2, \ldots, r-1;$$

$$\sigma^{(i)} = -\beta_i N_{i,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \operatorname{sign}(\phi_{i-1,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}));$$

$$(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)}) \notin \Sigma_{i-1}\},$$

where $i = 1, \ldots, r-1, \gamma_0$ contains only the origin $\sigma = \dot{\sigma} = \cdots \sigma^{(r-1)} = 0$. It is easy to check that

$$\Gamma_i = \bigcup_{j=0}^{i} \gamma_j, i = 1, \ldots, r-1$$

Lemma 3a: *Let β_j be chosen as in Lemma 3. There is such a vicinity O of the origin $\sigma = \dot{\sigma} = \cdots = \sigma^{(i)} = 0$ and such $\varepsilon > 0$ that each ε-swollen differential equation $\phi_{i,r} = 0, i = 1, \ldots, r-1$, provides for finite-time attraction of all trajectories from O into the ε-vicinity of Γ_{i-1}.*

Proof: Obviously, the ε-vicinity of Γ_i divides the space $\sigma, \dot{\sigma}, \ldots, \sigma^{(i)}$ in two parts (Lemma 1). Also 1-sliding mode on $\sigma^{(i-1)} = \xi_i(\sigma, \dot{\sigma}, \ldots, \sigma^{(i-1)})$ is ensured at continuity points of ξ_i out of the ε-vicinity of Γ_{t-1}. $\square$

Lemma 4: *With sufficiently large α any trajectory of the inclusion (30), (10) which leaves Γ at time t_0 returns in finite time to Γ at some time t_1. The time $t_1 - t_0$ and the maximal coordinate deviations from the initial point at the moment t_0 during $t_1 - t_0$ satisfy inequalities of the form*

$$t_1 - t_0 \leq c_0 \alpha^{-1} N_{r-1,r}(\sigma(t_0), \sigma(t_0), \dot{\sigma}(t_0), \ldots, \sigma^{(r-2)}(t_0))$$

$$|\Delta \sigma_i^{(r-j)}| \leq c_j \alpha^{-1} N_{r-1,r}(\sigma(t_0), \sigma(t_0), \dot{\sigma}(t_0), \ldots, \sigma^{(r-2)}(t_0))^j,$$

$$j = 1, 2, \ldots, r$$

where c_j are some positive constants.

Proof: Considering a small vicinity of the origin, we found that the first inequality is true when the initial conditions belong to a set of the form $N_{r-1,r} = \text{const}$. Thus, according to the homogeneity reasoning, it is true everywhere. Other inequalities are results of successive integration and the homogeneity reasoning with the same assumption $N_{r-1,r} = \text{const}$ with $t = t_0$. $\square$

The following lemma is a simple consequence of Lemma 4. It is illustrated by figure 2.

Lemma 4a: *With sufficiently large α any trajectory of the inclusion (30), (10) transfers in finite time into some vicinity of Γ and stays there. That vicinity contracts to Γ with $\alpha \to \infty$ uniformly within any compact area.*

Let

$$M(\eta) = \{\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)} | N_{r-1,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}) \leq \eta\}.$$

Lemma 5: *There is such $\eta_1 > 1$ that with sufficiently large α any trajectory of the inclusion (30), (10) starting from $M(1)$ never leaves $M(\eta_1)$.*

Lemma 5 is an obvious consequence of Lemmas 3a and 4a.

Lemma 6: *There are such $\eta_2 < 1$ and $T > 0$ that with sufficiently large α any trajectory of inclusion (30), (10) starting from $M(1)$ enters $M(\eta_2)$ within time T and stays in it.*

Proof: As follows from Lemmas 3a, 4a and 5, there are such successively embedded sufficiently small vicinities $O_i(\Gamma_i)$ of $\Gamma_i, i = 0, 1, \ldots, r-1, \Gamma_{r-1} = \Gamma$, that with sufficiently large α any trajectory of inclusion (30), (10) hits one of these vicinities in finite time and then transfers from one set to another in finite time according to the diagram:

$$O(\Gamma) = O_{r-1}(\Gamma_{r-1}) \to O_{r-2}(\Gamma_{r-2}) \to \cdots \to O_0(\Gamma_0)$$

$$= O_0(0) \quad \square$$

The invariance of inclusion (10), (30) with respect to

$$G_\nu: (t, \sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}) \mapsto (\nu t, \nu^r \sigma, \nu^{r-1} \dot{\sigma}, \ldots, \nu \sigma^{(r-1)})$$

implies that if $M(1)$ transfers in time T into $M(\eta_2)$, $M(\eta_2)$ transfers in time $\eta_2 T$ into $M(\eta_2^2)$, $M(\eta_2^2)$ transfers

in time $\eta_2^2 T$ into $M(\eta_2^3)$, etc. Thus, all trajectories starting from $M(1)$ enter the origin during the time $T \sum_{j=1}^{\infty} \eta^j < \infty$ and stay there. Due to the homogeneity reasoning the same is true with respect to initial conditions taken from a set $M(c)$ for any $c > 0$. Theorem 3 is proved. $\qquad\square$

Proof of Theorem 4: Similarly to the proof of Theorem 3 consider inclusion (30), (11) instead of (3), (11). According to the Taylor formula $\Delta\sigma_i^{(r-2)} = \tau\sigma_{i-1}^{(r-1)} + 0.5\tau^2\sigma^{(r)}(\theta),\ \theta \in [t_{i-1}, t_i]$. Hence, in its turn, inclusion (30), (11) may be replaced by the

$$\sigma^{(r)} \in [-C, C] - \alpha[K_{\mathrm{m}}, K_{\mathrm{M}}] \operatorname{sign}(\sigma^{(r-1)}$$

$$+\beta_{r-1} N_{r-1,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(r-2)}) \operatorname{sign}(\phi_{r-2,r}(\sigma, \dot{\sigma}, \ldots, \sigma^{(r-2)})$$

$$+0.5\tau[-C - \alpha K_{\mathrm{M}}, C + \alpha K_{\mathrm{M}}])$$

For any $\varepsilon > 0$ the inequality $\tau(C + \alpha K_{\mathrm{M}}) < \varepsilon N_{r-1,r}$ is satisfied outside of some bounded vicinity of the origin. Thus, similarly to the Theorem 3 proof, sufficiently small ε being taken, finite-time convergence into some vicinity D of the origin is provided. Obviously, the transformation $G_\nu: (t, \sigma, \dot{\sigma}, \ldots, \sigma^{(r-1)}) \mapsto (\nu t, \nu^r \sigma, \nu^{r-1}\dot{\sigma}, \ldots, \nu\sigma^{(r-1)})$ transfers (30), (11) into the same inclusion, but with the new measurement interval $\nu\tau$. Thus, $G_\nu D$ is an attracting set corresponding to that new value of the measurement interval. $\qquad\square$

Proof of Proposition 3: As is seen from the proof of Theorems 3 and 4 the only restriction on the choice of β_1 is formulated in Lemma 3, and α is chosen with respect to Lemma 3 and is sufficiently enlarged afterwards. Thus, at first the existence of the virtual sliding modes $\phi_{i,r} = 0,\ i = 1, \ldots, r-1$ is provided and than α is taken so large that the real motion will take place in arbitrarily small vicinity of these modes. Thus, the parameters may be chosen sufficiently large in the order $\beta_1, \ldots, \beta_{r-1},\ \alpha$ so that the convergence from any fixed compact region of initial conditions be arbitrarily fast, and at the same time the overregulation be arbitrarily small. $\qquad\square$

Appendix 2. Proofs of Theorems 5–7

Consider for simplicity differentiator (16). The proof for differentiator (15) is very similar. Introduce functions $\sigma_0 = z_0 - f_0(t),\ \sigma_1 = z_1 - \dot{f}_0(t), \ldots, \sigma_n = z_n - f_0^{(n)}(t),\ \xi = f(t) - f_0(t)$. Then any solution of (16) satisfies the following differential inclusion understood in the Filippov sense

$$\dot{\sigma}_0 = -\lambda_0 |\sigma_0 + \xi(t)|^{(n/(n+1))} \operatorname{sign}(\sigma_0 + \xi(t)) + \sigma_1 \quad (31)$$

$$\left.\begin{array}{l} \dot{\sigma}_1 = -\lambda_1 |\sigma_1 - \dot{\sigma}_0|^{(n-1)/n} \operatorname{sign}(\sigma_1 - \dot{\sigma}_0) + \sigma_2 \\[2mm] \vdots \\[2mm] \dot{\sigma}_{n-1} = -\lambda_{n-1} |\sigma_{n-1} - \dot{\sigma}_{n-2}|^{1/2} \operatorname{sign}(\sigma_{n-1} - \dot{\sigma}_{n-2}) + \sigma_n, \\[2mm] \dot{\sigma}_n \in -\lambda_n \operatorname{sign}(\sigma_n - \dot{\sigma}_{n-1}) + [-L, L] \end{array}\right\}$$

$$(32)$$

where $\xi(t) \in [-\varepsilon, \varepsilon]$ is a Lebesgue-measurable noise function. It is important to mark that (31), (32) does not 'remember' anything on the unknown input basic signal $f_0(t)$.

System (31), (32) is homogeneous with $\varepsilon = 0$, its trajectories are invariant with respect to the transformation

$$G_\eta: (t, \sigma_i, \xi, \varepsilon) \mapsto (\eta t, \eta^{n-i+1}\sigma_i, \eta^{n+1}\xi, \eta^{n+1}\varepsilon) \quad (33)$$

Define the main features of differential inclusion (31), (32) which hold with a proper choice of the parameters λ_i.

Lemma 7: *Let $\xi_\delta(t)$ satisfy the condition that the integral $\int |\xi_\delta(t)|\mathrm{d}t$ over a time interval δ is less than some fixed $K > 0$. Then for any $0 < S_i < S_i', i = 0, \ldots, n$, each trajectory of (31), (32) starting from the region $|\sigma_i| \le S_i$ does not leave the region $|\sigma_i| \le S_i'$ during this time interval if δ is sufficiently small.*

Lemma 8: *For each set of numbers $S_i > 0, i = 0, \ldots, n$, there exist such numbers $\Sigma_i > S_i, k_i > 0$ and $T > 0, \varepsilon_{\mathrm{M}} \ge 0$ that for any $\xi(t) \in [-\varepsilon, \varepsilon], \varepsilon \le \varepsilon_{\mathrm{M}}$, any trajectory of (31), (32) starting from the region $|\sigma_i| \le S_i$ enters within the time T, and without leaving the region $|\sigma_i| \le \Sigma_i$, the region $|\sigma_i| \le k_i \varepsilon^{n-i+1}$ and stays there forever.*

Mark that $-\dot{\sigma}_0$ plays rule of the disturbance ξ for the $(n-1)$th-order system (32). With $n = 0$ (31), (32) is reduced to $\dot{\sigma}_0 \in -\lambda_0 \operatorname{sign}(\sigma_0 + \xi(t)) + [-L, L]$, and Lemmas 7 and 8 are obviously true with $n = 0, \lambda_0 > L$. The lemmas are proved by induction. Let their statements be true for the system of the order $n-1$ with some choice of parameters $\bar{\lambda}_i,\ i = 0, \ldots, n-1$. Prove their statements for the nth order system (31), (32) with sufficiently large λ_0, and $\lambda_i = \bar{\lambda}_{i-1}, i = 1, \ldots, n$.

Proof of Lemma 7: Choose some $S_{Mi}, S_i < S_i' < S_{Mi},\ i = 0, \ldots, n$. Then

$$|\dot{\sigma}_0| \leq \lambda_0 \|\xi(t)\| + |\sigma_0|^{n/(n+1)} + |\sigma_1|$$

$$\leq \lambda_0 |\xi(t)|^{n/(n+1)} + \lambda_0 S_{M0}^{n/(n+1)} + S_{M1}$$

Thus, according to the Hölder inequality

$$\int |\dot{\sigma}_0| \mathrm{d}t \leq \lambda_0 \delta^{1/(n+1)} \left(\int |\xi(t)| \mathrm{d}t \right)^{n/(n+1)} + \delta(\lambda_0 S_{M0}^{n/(n+1)} + S_{M1})$$

Hence, $|\sigma_0| \leq S_0'$ with small δ. On the other hand $\dot{\sigma}_0$ serves as the input disturbance for the $(n-1)$th order system (32) and satisfies the conditions of Lemma 7, thus due to the induction assumption $|\sigma_i| \leq S_i'$, $i = 1, \ldots, n$ with small δ. $\square$

Lemma 9: *If for some $S_i < S_i', i = 0, \ldots, n, \xi \equiv 0$ and $T > 0$ any trajectory of (31), (32) starting from the region $|\sigma_i| \leq S_i'$ enters within the time T the region $|\sigma_i| \leq S_i$ and stays there forever, then the system (31), (32) is finite-time stable with $\xi \equiv 0$.*

Lemma 9 is a simple consequence of the invariance of (31), (32) with respect to transformation (33). The convergence time is estimated as a sum of a geometric series.

Proof of Lemma 8: Consider first the case $\varepsilon_M = 0$, i.e. $\xi = 0$. Choose some larger region $|\sigma_i| \leq S_i', S_i < S_i'$, $i = 0, \ldots, n$, and let $\Sigma_i > S_i', i = 0, \ldots, n-1$, be some upper bounds chosen with respect to Lemma 8 for (32) and that region. It is easy to check that for any $q > 1$ with sufficiently large λ_0 the trajectory enters the region $|\sigma_0| \leq (q\Sigma_1/\lambda_0)^{(n+1)/n}$ in arbitrarily small time. During that time $\dot{\sigma}_0$ does not change its sign. Therefore, $\int |\dot{\sigma}_0| \mathrm{d}t \leq S_0'$ for sufficiently large λ_0. Thus, the disturbance $-\dot{\sigma}_0$ entering subsystem (32) satisfies Lemma 7 and the inequalities $|\sigma_i| \leq S_i'$ are kept with $i \geq 1$. As follows from (31), from that moment on $|\dot{\sigma}_0| \leq 3\Sigma_1$ is kept with a properly chosen q.

Differentiating (31) with $\xi = 0$ achieve

$$\ddot{\sigma}_0 = -\lambda_0 |\sigma_0|^{-1/(n+1)} \dot{\sigma}_0 + \dot{\sigma}_1$$

where according to (32) the inequality $|\dot{\sigma}_1| \leq 4\lambda_1 \Sigma_1 + \Sigma_2$ holds. Thus, with sufficiently large λ_0 within arbitrarily small time the inequality $|\dot{\sigma}_0| \leq \lambda_0^{-1}(4\lambda_1\Sigma_1 + \Sigma_2)(q\Sigma_1/\lambda_0)^{1/n}$ is established. Its right-hand side may be made arbitrarily small with large λ_0. Thus, due to Lemma 8 for $(n-1)$-order system (32) and to Lemma 9, the statement of Lemma 8 is proved with $\varepsilon_M = 0$.

Let now $\varepsilon_M > 0$. As follows from the continuous dependence of the solutions of a differential inclusion on the right-hand side (Filippov 1988) with sufficiently small ε_M all trajectories concentrate within a finite time in a small vicinity of the origin. The asymptotic features of that vicinity with ε_M changing follow now from the homogeneity of the system. $\square$

Theorems 5 and 6 are simple consequences of Lemma 8 and the homogeneity of the system. To prove Theorem 7 it is sufficient to consider

$$\dot{\sigma}_0 = -\lambda_0 |\sigma_0(t_j)|^{n/(n+1)} \operatorname{sign}(\sigma_0(t)) + \sigma_1 \qquad (34)$$

instead of (31) with $t_j \leq t < t_{j+1} = t_j + \tau$. The resulting hybrid system (34), (32) is invariant with respect to the transformation

$$(t, \tau, \sigma_i) \mapsto (\eta t, \eta \tau, \eta^{n-i+1} \sigma_i) \qquad (35)$$

On the other hand, it may be considered as system (31), (32), $|\xi| \leq \varepsilon$, with arbitrarily small ε for any fixed region of the initial values when τ is sufficiently small. Applying successively Lemma 8 and the homogeneity reasoning with transformation (35) achieve Theorem 7. $\square$

Remark: It is easy to see that the above proof may be transformed in order to obtain a constructive upper estimation of the convergence time. Also the accuracy may be estimated by means of the suggested inductive approach.

References

ATASSI, A. N., and KHALIL, H. K., 2000, Separation results for the stabilization of nonlinear systems using different high-gain observer designs. *Systems and Control Letters*, **39**, 183–191.

BARTOLINI, G., FERRARA, A., LEVANT, A., and USAI, E., 1999 a, On second order sliding mode controllers. In K. D. Young and U. Ozguner (Eds) *Variable Structure Systems, Sliding Mode and Nonlinear Control* (*Lecture Notes in Control and Information Science*, 247) (London: Springer-Verlag), pp. 329–350.

BARTOLINI, G., FERRARA, A., and USAI, E., 1998, Chattering avoidance by second-order sliding mode control. *IEEE Transactions on Automatic Control*, **43**, 241–246.

BARTOLINI, G., FERRARA, A., and USAI, E., 1999 b, On multi-input second order sliding mode control of nonlinear systems with uncertainty. *Proceedings of the 38th Conference on Decision and Control*, Phoenix, AZ, USA, December 1999.

BARTOLINI, G., PISANO, A., and USAI, E., 2000, First and second derivative estimation by sliding mode technique. *Journal of Signal Processing*, **4**, 167–176.

DABROOM, A., and KHALIL, H. K., 1997, Numerical differentiation using high-gain observers. *Proceedings of the 37th Conference on Decision and Control (CDC'97)*, San Diego, CA, USA, December 1997, pp. 4790–4795.

ELMALI, H., and OLGAC, N., 1992, Robust output tracking control of nonlinear MIMO systems via sliding mode technique. *Automatica*, **28**, 145–151.

EMELYANOV, S. V., KOROVIN, S. K., and LEVANTOVSKY, L. V., 1986, Higher order sliding regimes in the binary control systems. *Soviet Physics, Doklady*, **31**, 291–293.

FILIPPOV, A. F., 1988, *Differential Equations with Discontinuous Right-Hand Side* (Dordrecht, The Netherlands: Kluwer).

FRIDMAN, L. M., 1990, Singular extension of the definition of discontinuous systems and stability. *Differential Equations*, **10**, 1307–1312.

FRIDMAN, L., and LEVANT, A., 1996, Sliding modes of higher order as a natural phenomenon in control theory. In F. Garofalo and L. Glielmo (Eds) *Robust Control via Variable Structure and Lyapunov Techniques* (*Lecture Notes in Control and Information Science*, 217) (London: Springer-Verlag), pp. 107–133.

FRIDMAN, L., and LEVANT, A., 2002, Higher order sliding modes. In W. Perruquetti and J. P. Barbot (Eds) *Sliding Mode Control in Engineering* (New York: Marcel Dekker), pp. 53–101.

FURUTA, K., and PAN, Y., 2000, Variable structure control with sliding sector. *Automatica*, **36**, 211–228.

GOLEMBO, B. Z., EMELYANOV, S. V., UTKIN, V. I., and SHUBLADZE, A. M., 1976, Application of piecewise-continuous dynamic systems to filtering problems, *Automation and Remote Control*, **37**, Part 1, 369–377.

ISIDORI, A., 1989, *Nonlinear Control Systems*, 2nd edn (New York: Springer Verlag).

ISIDORI, A., TEELC, A. R., and PRALY, L., 2000, A note on the problem of semiglobal practical stabilization of uncertain nonlinear systems via dynamic output feedback. *Systems and Control Letters*, **39**, 165–171.

KRUPP, D., SHKOLNIKOV, I. A., and SHTESSEL, Y. B., 2001, 2-sliding mode control for nonlinear plants with parametric and dynamic uncertainties. *Proceedings of AIAA Guidance, Navigation, and Control Conference*, Denver, CO, USA, AIAA paper No. 2000–3965.

KOKOTOVIC, P., and ARCAK, M., 2001, Constructive nonlinear control: a historical perspective. *Automatica*, **37**, 637–662.

LANDAU, I. D., LOZANO, R., and M'SAAD, M., 1998, *Adaptive Control* (London: Springer-Verlag).

LEVANT, A. (Levantovsky, L.V.), 1993, Sliding order and sliding accuracy in sliding mode control. *International Journal of Control*, **58**, 1247–1263.

LEVANT, A., 1998 a, Robust exact differentiation via sliding mode technique. *Automatica*, **34**, 379–384.

LEVANT, A., 1998 b, Arbitrary-order sliding modes with finite time convergence. *Proceedings of the 6th IEEE Mediterranean Conference on Control and Systems*, June 9–11, Alghero, Sardinia, Italy.

LEVANT, A., 1999, Controlling output variables via higher order sliding modes. *Proceedings of the European Control Conference*, September, Karlsruhe, Germany.

LEVANT, A., 2001 a, Universal SISO sliding-mode controllers with finite-time convergence. *IEEE Transactions on Automatic Control*, **46**, 1447–1451.

LEVANT, A., 2001 b, Higher order sliding modes and arbitrary-order exact robust differentiation. *Proceedings of the European Control Conference*, Porto, Portugal, pp. 996–1001.

LEVANT, A., 2002, Universal SISO output-feedback controller. *Proceedings of IFAC'2002*, Barcelona, Spain, July 21–26.

LEVANTOVSKY, L. V., 1985, Second order sliding algorithms. Their realization. In *Dynamics of Heterogeneous Systems* (Moscow: Institute for System Studies), pp. 32–43 (in Russian).

MAN, Z., PAPLINSKI, A. P., and WU, H. R., 1994, A robust MIMO terminal sliding mode control for rigid robotic manipulators. *IEEE Transactions on Automatic Control*, **39**, 2464–2468.

MURRAY, R., and SASTRY, S., 1993, Nonholonomic motion planning: steering using sinusoids. *IEEE Transactions on Automatic Control*, **38**, 700–716.

RIOS-BOLÍVAR, M., ZINOBER, A. S. I., and SIRA-RAMÍREZ, H., 1997, Dynamical adaptive sliding mode output tracking control of a class of nonlinear systems. *International Journal of Robust and Nonlinear Control*, **7**, 387–405.

ROSIER, L., 1992, Homogeneous Lyapunov function for homogeneous continuous vector field. *Systems Control Letters*, **19**, 467–473.

SIRA-RAMÍREZ, H., 1993, On the dynamical sliding mode control of nonlinear systems. *International Journal of Control*, **57**, 1039–1061.

SLOTINE, J.-J. E., and LI, W., 1991, *Applied Nonlinear Control* (London: Prentice-Hall).

SPURGEON, S. K., and LU, X. Y., 1997, Output tracking using dynamic sliding mode techniques. *International Journal of Robust and Nonlinear Control, Special Issue on New Trends in Sliding Mode Control*, **7**, 407–427.

SU, W.-C., DRAKUNOV, S. V., and ÖZGÜNER, Ü., 1994, Implementation of variable structure control for sampled-data systems. *Proceedings of IEEE Workshop on Robust Control via Variable Structure and Lyapunov Techniques*, Benevento, Italy, pp. 166–173.

UTKIN, V. I., 1992, *Sliding Modes in Optimization and Control Problems* (New York: Springer Verlag).

WU, Y., YU, X., and MAN, Z., 1998, Terminal sliding mode control design for uncertain dynamic systems. *Systems and Control Letters*, **34**, 281–287.

YU, X., and XU, J. X., 1996, Nonlinear derivative estimator. *Electronic Letters*, **32**, 1445–1447.