Levant A. (1998). Robust exact differentiation via sliding mode technique. *Automatica*, 34(3), 379-384

Robust exact differentiation via sliding mode technique

Arie Levant *

Abstract. The main problem in differentiator design is to combine differentiation exactness with robustness in respect to possible measurement errors and input noises. The proposed differentiator provides for proportionality of the maximal differentiation error to the square root of the maximal deviation of the measured input signal from the base signal. Such an order of the differentiation error is shown to be the best possible when the only information known on the base signal is an upper bound for Lipschitz's constant of the derivative.

Key Words: Differentiators; measurement noise; nonlinear control; sliding mode; robustness.

Introduction

Differentiation of signals given in real time is an old and well-known problem. In many cases construction of a special differentiator may be avoided. For example, if the signal satisfies a certain differential equation or is an output of some known dynamic system, the derivative of the given signal may be calculated as a derivative with respect to some known dynamic system. Thus, the problem is reduced to the well-known observation and filtration problems. In other cases construction of a differentiator is inevitable. However, the ideal differentiator could not be realized. Indeed, together with the basic signal it would also have to differentiate any small high-frequency noise which always exists and may have a large derivative.

Constructing a differentiator as a single unit is a traditional problem for signal processing theory (Pei and Shyu 1989, Kumar and Roy 1988, Rabiner and Steiglitz 1970). The main approach is to approximate the transfer function of the ideal

* Dr. Arie Levant.

Institute for Industrial Mathematics, 22 Ha-Histadrut St., Beer-Sheva 84213, Israel

Tel: 972-7-6231313 Fax: 972-7-6231211 E-mail: levant@bgumail.bgu.ac.il

differentiator by that of some linear dynamic system. The problem being solved mainly for definite frequency bands of the signal and the noise, low-pass filters are used to damp noises. Stochastic features of the signal and the noise may also be considered (Carlsson *et. al* 1991). In the latter case the stochastic models of both the signal and the noise are presumed to be known. In any case the resulting differentiator does not calculate exact derivatives of arbitrary noise-free signals even when the signal frequency band is bounded.

If nothing is known on the structure of the signal except some differential inequalities, then sliding modes (Utkin 1992) are used. In the absence of noise the exact information on the signal derivative may be obtained in that case by averaging high-frequency switching signals. Also sliding observers (Slotine *et al.* 1987) or observers with large gains (Nicosia *et. al* 1991) are successfully employed. However, in all these cases the exact differentiation is provided only when some differentiator parameters tend to inadmissible values (like infinity). Thus, here too the resulting differentiator cannot calculate exact derivatives of noise-free signals.

The performance of the known differentiators follows the following principle: only approximate differentiation is provided in the absence of noise, at the same time the differentiator is insensitive to any high-frequency signal components considered to be noises. Thus, differentiation is robust but not exact, the error does not tend to zero in the presence of vanishing noise at any fixed time, and no asymptotic error analysis is sensible for any fixed differentiator parameters and time.

Another principle employed here combines exact differentiation (with finite transient time) for a large class of inputs with robustness in respect to small noises of any frequency. A known approach (Golembo *et al.* 1976) is chosen: high-quality tracking of $f(t)$ by $x(t)$, $\dot{x} = u$, having been provided, control $u(t)$ may be used for evaluation of $\dot{f}(t)$. The new result is attained here due to application of a 2-sliding algorithm (Levantovsky 1985, Emelyanov *et al.* 1986, Levant (Levantovsky) 1993, Fridman and Levant 1996) which forms continuous control $u(t)$ providing for keeping the equalities $\sigma = x - f(t) = 0$, $\dot{\sigma} = u - \dot{f}(t) = 0$ after a finite-time transient process. The purposes of this paper are

- to clear some inherent restrictions on exact robust differentiation and its error asymptotics;
- to propose a robust first order differentiator that is exact on signals with a given upper bound for Lipschitz's constant of the derivative;
- to ensure the best possible error asymptotics order when the input noise is a measurable (Lebesgue) bounded function of time.

Robust exact differentiation limitations

Let input signals belong to the space $M[a,b]$ of measurable functions bounded on a segment $[a,b]$ and let $||f|| = \sup|f(t)|$. Define *abstract differentiator* as a map associating an output signal with any input signal. A differentiator is called *exact* on some input if the output coincides with its derivative. The differentiator order is the order of the derivative which it produces. Differentiator D is called *robust* on some input $f(t)$ if the output tends uniformly to $Df(t)$ while the input signal tends uniformly to $f(t)$. A differentiator is called *correct* on some input if it is exact and robust on it. The ideal differentiator cannot be considered as an abstract one, for it does not operate for nondifferentiable inputs.

Being exact on two inputs, any differentiator will actually differentiate a difference between these inputs which may be considered as a noise. Thus, differentiator design is a simple trade-off: the denser the exactness class in $M[a,b]$, the more sensitive will the differentiator be to small noises. For example, being correct on a thin class of constant inputs, the differentiator producing identical zero is totally insensitive to noises. It is easily seen that the correctness set of any abstract differentiator cannot be locally dense in the set of continuous functions, otherwise small noises with large derivatives would be exactly differentiated, which contradicts robustness. In particular, no differentiator is correct on all smooth functions or on all polynomials.

Let $W(C,n)$ be the set of all input signals whose $(n$-1)-th derivatives have Lipschitz's constant $C > 0$. The statements below are valid for any sufficiently small $\varepsilon > 0$ and noises not exceeding ε in absolute value.

Proposition 1. *No exact on $W(C,n)$, $n>0$, differentiator of the order $i \leq n$ may provide for accuracy better than $C^{i/n}\varepsilon^{(n-i)/n}$.*

Proof. Consider noise $\nu(t) = \varepsilon \sin(C/\varepsilon)^{1/n}t$. It is easy to check that $\sup|\nu^{(i)}(t)| = C^{i/n}\varepsilon^{(n-i)/n}$ for $i = 0,1, ..., n$ and small ε. The differentiator has to differentiate properly both 0 and $0 + \nu(t)$. □

Proposition 2. *There are such constants $\gamma_i(n) \geq 1$, $n = 1,2, ...,$ $i = 0, 1, ..., n$, $\gamma_0(n) = 1$, $\gamma_n(n) = 2$, that for any $i \leq n$ there exists an i-th order differentiator D, correct on any $f \in W(C,n)$, which provides for the accuracy $||D(f+\nu)-f^{(i)}|| \leq \gamma_i(n)\, C^{i/n}\varepsilon^{(n-i)/n}$ for any noise ν, $||\nu|| < \varepsilon$.*

Proof. Let Δ be a map associating with any $g \in M[a,b]$ some function Δg closest to g in $W(C,n)$. Such a function exists, for any bounded subset of $W(C,n)$ is precompact, but it is certainly not unique. Define $Dg = (\Delta g)^{(i)}$. The Proposition is now a simple consequence of Lemma 1. □

Lemma 1. *There are such constants $\beta_i(n) \geq 1$, $i = 0,2, ..., n$, $\beta_0(n) = \beta_n(n) = 1$, that the following inequalities hold:*

$$C^{i/n}\varepsilon^{(n-i)/n} \leq \sup_{f\in W(C,n),\ \sup|f(t)|\leq\varepsilon}\ \sup_t |f^{(i)}(t)| \leq \beta_i(n)\, C^{i/n}\varepsilon^{(n-i)/n}.$$

Here and below all proofs are in Appendixes. These are some evaluations of $\beta_i(n)$: $\beta_1(2)=2\sqrt{2}$, $\beta_1(3)=7.07$, $\beta_2(3)=6.24$.

Remarks.

- With C=0 $W(C,n)$ is the set of polynomials of the degree n-1. In that case Proposition 1 is trivial, Proposition 2 and Lemma 1 are not valid. The restrictions discussed here do not preclude possibilities to receive better differentiation accuracy, provided another exactness set is considered. For instance, it may be shown that the differentiator suggested in the proof of Proposition 2 provides for the accuracy linear on ε in the case C=0. The corresponding gain tends to infinity when $n\to\infty$ or the segment length tends to 0. $W(0,n)$ being too thin for reasonable n, such differentiators are suitable for local usage only.

- All above reasonings are also true in the case when the differentiators are allowed to have transient time uniformly bounded by a constant less than the segment length.
- The above results also hold for infinite time intervals, the inputs being measurable locally bounded functions and ε being *any* positive number. Only a minor change in the proof of Proposition 2 is needed.

Practical first-order robust exact differentiator

The above abstract differentiators were not intended for realization. Consider now a practical real-time differentiation problem. Let input signal $f(t)$ be a measurable locally bounded function defined on $[0, \infty)$ and let it consist of a base signal having a derivative with Lipschitz's constant $C > 0$ and a noise. In order to differentiate the unknown base signal, consider the auxiliary equation

$$\dot{x} = u. \tag{3}$$

Applying a modified 2-sliding algorithm (Levant 1993) to keep $x - f(t) = 0$, obtain

$$u = u_1 - \lambda\, |x - f(t)|^{1/2} \operatorname{sign}(x - f(t)), \quad \dot{u}_1 = -\alpha \operatorname{sign}(x - f(t)), \tag{4}$$

where $\alpha, \lambda > 0$. Here $u(t)$ is the output of the differentiator. Solutions of system (3), (4) are understood in the Filippov sense (Filippov 1988).

Define a function $\Phi(\alpha,\lambda,C) = |\Psi(t_*)|$, where $(\Sigma(t), \Psi(t))$ is the solution of

$$\dot{\Sigma} = -|\Sigma|^{\frac{1}{2}} + \Psi, \qquad \dot{\Psi} = \begin{cases} -\frac{1}{\lambda^2}(\alpha - C), & -|\Sigma|^{\frac{1}{2}} + \Psi > 0, \\ -\frac{1}{\lambda^2}(\alpha + C), & -|\Sigma|^{\frac{1}{2}} + \Psi \le 0, \end{cases} \qquad \Sigma(0)=0,\ \Psi(0)=1, \tag{5}$$

$\alpha > C$, $\lambda \neq 0$, and $t_* = \inf\{t \mid t>0,\ \Sigma(t)=0,\ \Psi(t)<0\}$. It is easy to check that $t_* < \infty$. In practice $\Phi(\alpha,\lambda,C)$ is to be calculated by computer simulation.

Theorem 1 (convergence criterion). *Let* $\alpha > C > 0$, $\lambda > 0$, $\Phi(\alpha,\lambda,C) < 1$. *Then, provided* $f(t)$ *has a derivative with Lipschitz's constant* C $(f \in W(C,2))$, *the equality* $u(t) = \dot{f}(t)$ *is fulfilled identically after a finite time transient process. There is no convergence of* $u(t)$ *to* $\dot{f}(t)$ *for some* $f \in W(C,2)$ *if* $\Phi(\alpha,\lambda,C) > 1$.

The less $\Phi(\alpha,\lambda,C)$, the faster the convergence. Φ is obviously the same for all (α,λ,C) with $\alpha = \mu_1 C$, $\lambda = \mu_2 C^{\frac{1}{2}}$ where $\mu_1 > 1$, $\mu_2 > 0$ are some constants. Also, any increase of λ decreases Φ.

Following are the *sufficient conditions* for the convergence of $u(t)$ to $\dot{f}(t)$:

$$\alpha > C,\ \ \lambda^2 \geq 4C\frac{\alpha+C}{\alpha-C}. \tag{6}$$

Condition (6) results from a very crude estimation. Calculation of Φ shows that many other values, *e.g.* $\lambda = C^{\frac{1}{2}}$, $\alpha = 1.1C$ (Φ =0.988), or $\lambda = 0.5C^{\frac{1}{2}}$, $\alpha = 4C$ (Φ =0.736), may also be taken.

The conditions $\alpha > C > 0$, $\lambda > 0$, $\Phi < 1$ are assumed to be satisfied in the following Theorems.

Theorem 2. *Let input signal be presented in the form* $f(t) = f_0(t) + v(t)$, *where* $f_0(t)$ *is a differentiable base signal,* $f_0(t)$ *has a derivative with Lipschitz's constant* $C > 0$, *and* $v(t)$ *is a noise,* $|v(t)| \leq \varepsilon$. *Then there exists such a constant* $b > 0$ *dependent on* $\frac{\alpha - C}{\lambda^2}$ *and* $\frac{\alpha + C}{\lambda^2}$ *that after a finite time the inequality* $|u(t)-\dot{f}_0(t)| < \lambda b\varepsilon^{\frac{1}{2}}$ *holds.*

If λ *and* α *are chosen in the form* $\alpha = \mu_1 C$, $\lambda = \mu_2 C^{\frac{1}{2}}$, *then the inequality* $|u(t)-\dot{f}_0(t)| < \tilde{b}\, C^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}$ *holds for some* $\tilde{b}\,(\mu_1,\mu_2) > 0$.

Let f, x, u_1 be measured at discrete times with time interval τ, and let t_i, t_{i+1}, t be successive measurement times and the current time, $t \in [t_i, t_{i+1})$. As a result, achieve the following modified algorithm:

$$\dot{x} = u, \tag{7}$$

$$u = u_1(t_i) - \lambda\, |x(t_i) - f(t_i)|^{\frac{1}{2}} \operatorname{sign}(x(t_i) - f(t_i)),\ \ \dot{u}_1 = -\alpha \operatorname{sign}(x(t_i) - f(t_i)). \tag{8}$$

Theorem 3. *Provided* $f(t)$ *has a derivative with Lipschitz's constant* $C > 0$, *algorithm* (7), (8) *enables the inequality* $|u(t)-\dot{f}(t)| < a\lambda^2\tau$ *to hold after a finite time transient process. Here* $a > 0$ *is some constant dependent on* $\frac{\alpha - C}{\lambda^2}$ *and* $\frac{\alpha + C}{\lambda^2}$.

If the discrete measurements are carried out with some small measurement noise of magnitude ε and $\tau << C^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}$ (or $C^{\frac{1}{2}}\varepsilon^{\frac{1}{2}} << \tau$), then an infinitesimal of a higher order has to be added to the right-hand side of the inequality in the statement of Theorem 2 (or Theorem 3). It may be shown that in the general case at least the accuracy of the order $(\tau+\varepsilon)^{1/2}$ is provided (very crude estimation).

The transient process time is uniformly bounded if the initial deviations $|x(t_0) - f(t_0)|$ and $|u(t_0)- \dot{f}(t_0)|$ are bounded. This may be attained by any inexact preliminary evaluation of $f(t_0)$ and $\dot{f}(t_0)$. In such a case the transient time may be arbitrarily shortened.

Computer simulation

It was taken that $t_0 = 0$, initial values of the internal variable $x(0)$ and the measured input signal $f(0)$ coincide, initial value of the output signal $u(0)$ is zero. The simulation was carried out by the Euler method with measurement and integration steps equaling 10^{-4}.

Compare the proposed differentiator (3), (4) with a simple linear differentiator described by the transfer function $p/(0.1p+1)^2$. Such a differentiator is actually a combination of the ideal differentiator and a low-pass filter. Let $\alpha = 8$, $\lambda = 6$. Checking the convergence criterion, achieve $\Phi = 0.064$ for $C = 2$ and $\Phi = 0.20$ for $C = 7$. The output signals for inputs $f(t) = \sin t + 5t$, $f(t) = \sin t + 5t + 0.01 \cos 10t$, and $f(t) = \sin t + 5t + 0.001 \cos 30t$ and ideal derivatives $\dot{f}(t)$ are shown in Fig. 1. The linear differentiator is seen not to differentiate exactly. At the same time it is highly insensitive to any signals with frequency above 30. The proposed differentiator handles properly any input signal f with $\ddot{f} \leq 7$ regardless the signal spectrum.

Let α=2.2, λ=2. Checking the convergence criterion achieve Φ=0.596 for C=2. The output for the input base signal $f(t) = \sin t +5t +1$ in the presence of a measurement high-frequency noise with magnitude 0.04 is shown in Fig. 2a. Simulation shows the differentiation accuracy to take on successively the values 0.426, 0.213, 0.106, 0.053, while the noise magnitude takes on the values 0.04, 0.01, 0.0025, $6.25\cdot10^{-4}$ respectively. Thus the differentiation error is proportional to the square root of the noise magnitude with the gain 2.13 (the same gain is 5.9 for $\alpha = 8$, $\lambda = 6$). With precise measurements the accuracy $1.2\cdot10^{-3}$ was achieved. Remind that the gain is not less than 1 (Proposition

1) and in the absence of noise the accuracy is proportional to the measurement time step (Theorem 3).

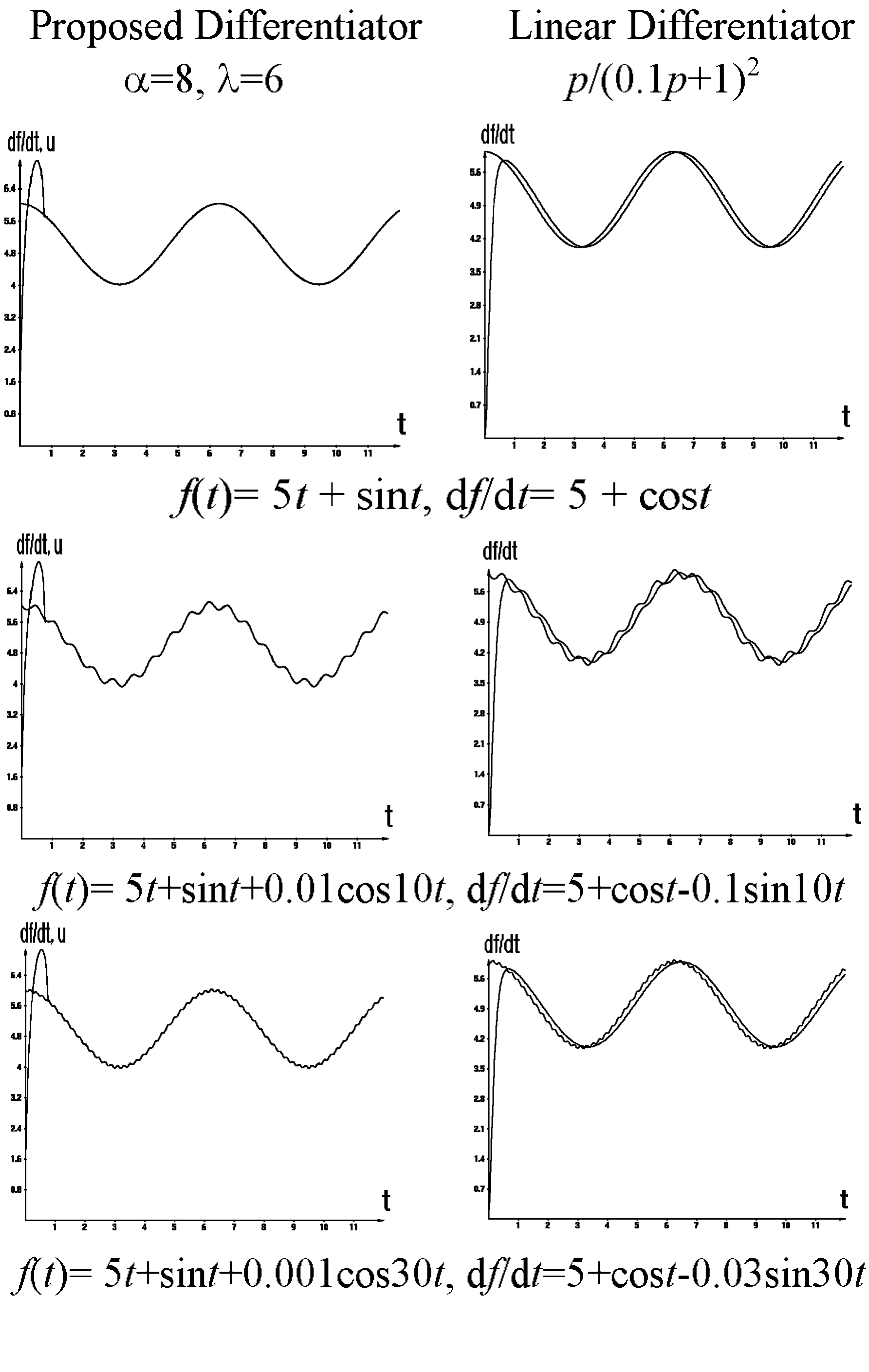
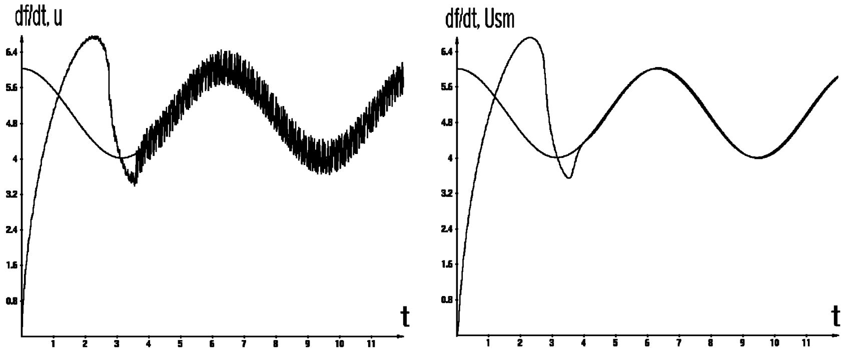


Fig. 1: Comparison of the proposed differentiator with a linear one.

The output signal of the differentiator consists, essentially, of the accurate derivative and some small high-frequency noise. Therefore, implementation of a smoothing element may be practically useful if significant noise is assumed. Output u_{sm}

of the simple smoothing element p/(0.05p+1) for the base signal $f(t)$ = sin t +5 t + 1 and noise magnitude 0.04 is given in Fig. 2b (the accuracy achieved is 0.054). However, the ideal differentiation ability in the absence of noises is inevitably lost in this case (the accuracy sup| $\dot{f}$ - u_{sm}| = 0.05 is achieved). Such a differentiator is insensitive to high-frequency components of the input signal.



a. Before smoothing b. After smoothing

Fig. 2: Mutual graphs of the differentiator output $u(t)$ (a) and smoothed output $u_{sm}(t)$ (b) with ideal derivative $\dot{f}(t)$ for the input signal $f(t)$ = sin t + 5 t + 1 in the presence of a noise with magnitude 0.04

Conclusions

Inherent restrictions on exact robust differentiation and its error asymptotics were found. The existence of an arbitrary-order robust differentiator with the optimal order of error asymptotics was established.

A first-order robust exact differentiator was proposed providing for maximal derivative error to be proportional to the square root of the input noise magnitude after a finite time transient process. This asymptotics was shown to be the best attainable in the case when the only restriction on the input signal is that Lipschitz's constant of its derivative is bounded by a given constant and the noise is a measurable bounded function of time.

Discrete measurements with a small time step were shown not to be destructive to the differentiator features. In the absence of noises the differentiation error is proportional to the measurement step.

The differentiator considered features simple form and easy design. It may be employed both in real-time control systems and for numeric differentiation. Its use is preferable in high precision systems with small noises. In the presence of considerable measurement noises a simple smoothing element may be implemented. However, this leads to loss of ideal differentiation in the absence of noises.

The differentiator allows, obviously, successive implementation for higher-order exact robust differentiation. However, the optimal error asymptotics will not be attained in that case.

References

B.Carlsson, A. Ahlen, M. Sternad (1991). Optimal differentiation based on stochastic signal models. *IEEE Trans. Signal Processing,* **39(2),** 341-353.

S.V. Emelyanov, S.K. Korovin, and L.V. Levantovsky (1986). Higher order sliding modes in the binary control systems. *Soviet Physics, Doklady,* **31(4)**, 291-293.

A.F. Filippov (1988) Differential Equations with Discontinuous Right-Hand Side. Kluwer, Dordrecht, the Netherlands.

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

B.Z. Golembo, S.V. Emelyanov, V.I. Utkin, and A.M. Shubladze (1976). Application of piecewise-continuous dynamic systems to filtering problems, *Automation and remote control,* **37(3)**, Part 1, 369-377.

B. Kumar and S.C.D. Roy (1988). Design of digital differentiators for low frequencies, *Proc. IEEE,* **76**, 287-289.

A. Levant (Levantovsky, L.V.) (1993). Sliding order and sliding accuracy in sliding mode control, *International Journal of Control,* **58(6),** 1247-1263.

L.V. Levantovsky (1985). Second order sliding algorithms: their realization, in: "Dynamics of Heterogeneous Systems", 32-43. Institute for System Studies, Moscow, [in Russian].

S. Nicosia, A. Tornambe, and P. Valigi (1991). A solution to the generalized problem of nonlinear map inversion, Systems & Control Letters, **17**, 383-394.

S.-C. Pei and J.-J. Shyu (1989). Design of FIR Hilbert Transformers and differentiators by eigenfilter, *IEEE Trans. Acoust., Speech, Signal Processing,* **ASSP(37)**, 505-511.

L.R. Rabiner and K. Steiglitz (1970). The design of wide-band recursive and nonrecursive digital differentiators, *IEEE Trans. Audio Electroacoust.*, **AU(18)**, 204-209.

J.-J.E Slotine, J.K. Hedrick and E.A. Misawa (1987). “On Sliding Observers for Nonlinear Systems”, J. of Dynamic Sys., Measurement, and Control, **109(9)**, 245-252.

V.I. Utkin (1992). Sliding Modes in Optimization and Control Problems. Springer-Verlag, New York.

Appendices

Proof of Lemma 1. The first inequality follows from the proof of Proposition 1. The Lagrange Theorem implies that there are constants L_k satisfying the property that for any k-smooth function $f(t)$, $\| f \| \le \varepsilon$, and any $T > 0$ there is a point t_* on any segment of the length $L_k T$ where $| f^{(k)}(t_*)| \le 2^k \varepsilon / T^k$. Obviously, $L_0 = 0$, $L_k = 2\, L_{k-1} + 1$.

Denote $S_i(\varepsilon,C,n) = \sup\limits_{f \in W(C,n),\ \sup|f(t)| \le \varepsilon} \ \sup\limits_{t} | f^{(i)}(t)|$. Using identities $S_n(\varepsilon,C,n)=C$, $S_0(\varepsilon,C,n) = \varepsilon$, the inequality $S_i(\varepsilon,C,n) \le S_i(\varepsilon, S_{n-1}(\varepsilon,C,n),n-1)$, and minimizing with respect to T the right-hand side of the inequality $| f^{(n-1)}(t)| \le 2^{n-1}\varepsilon/T^{n-1} + CL_{n-1}T$, obtain a recursive-on-n definition of $\beta_i(n)$

$$\beta_0(n) = \beta_n(n) = 1,$$

$$\beta_i(n) = \beta_i(n-1)\ \beta_{n-1}(n)^{i/(n-1)},\ i < n,$$

$$\beta_{n-1}(n) = 2^{\frac{n-1}{n}} L_{n-1}^{\frac{n-1}{n}} [\tfrac{1}{n-1}^{\frac{n-1}{n}} + (n-1)^{\frac{1}{n}}]. \quad \square$$

The plan of the Theorems proofs is as follows. Using the inclusion $\ddot{f} \in [-C,C]$, true almost everywhere, system (3), (4) is replaced by an autonomous differential inclusion on plane $\sigma\dot{\sigma}$, $\sigma = x\text{-}f(t)$. Intersections with axis $\dot{\sigma}$ having been studied, the paths are shown to make an infinite number of rotations around the origin converging to it in finite time. With a small measurement step and in the presence of small noises a small attraction set appears around the origin. After changing coordinates a linear transformation preserving the paths is found, which allows to evaluate the attraction set asymptotics.

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

Proof of Theorem 1. Let $\sigma = x\text{-}f(t)$. By calculating achieve

$$\ddot{\sigma} = -\ \ddot{f} - \tfrac{1}{2}\lambda\dot{\sigma}|\sigma|^{-\frac{1}{2}} - \alpha \operatorname{sign} \sigma. \tag{9}$$

It does not matter that $\ddot{f}$ exists almost everywhere, but not at any t, for solutions are understood in the Filippov sense. Strictly speaking, (9) is valid only for $\sigma \neq 0$. Nevertheless, as follows from (Filippov 1988), solutions of (3), (4) (and, therefore, also $\sigma(t) = x(t)\text{-}f(t)$ and $\dot{\sigma}(t) = u(t)\text{-}\dot{f}(t)$) are well defined for any initial conditions ($\sigma(t_0)$, $\dot{\sigma}(t_0)$) and any $t > t_0$.

Remind that $|\ddot{f}| \leq C$. Denote $R = \alpha + \ddot{f} \operatorname{sign} \sigma$. Then $R \in [\alpha - C, \alpha + C]$ and

$$\ddot{\sigma} = -\ \tfrac{1}{2}\lambda\dot{\sigma}|\sigma|^{-\frac{1}{2}} - R \operatorname{sign} \sigma. \tag{10}$$

Any solution of (9) has to satisfy the following differential inclusion understood in the Filippov sense:

$$\ddot{\sigma} \in -\ \tfrac{1}{2}\lambda\dot{\sigma}|\sigma|^{-\frac{1}{2}} - [\alpha - C, \alpha + C] \operatorname{sign} \sigma. \tag{11}$$

The operations on sets are naturally understood here as sets of operation results for all possible combinations of the operand set elements. Consider the trajectory of (11) on plane $\sigma\dot{\sigma}$. Let $\sigma = 0$, $\dot{\sigma} = \dot{\sigma}_0 > 0$ at the initial moment (Fig. 3). Any trajectory Γ of (11) lies between trajectories *I*, *II* of (10) given respectively by

$$R = \begin{cases} \alpha - C, & \sigma\dot{\sigma} > 0, \\ \alpha + C, & \sigma\dot{\sigma} \le 0 \end{cases}, \qquad R = \begin{cases} \alpha + C, & \sigma\dot{\sigma} > 0, \\ \alpha - C, & \sigma\dot{\sigma} \le 0. \end{cases}$$

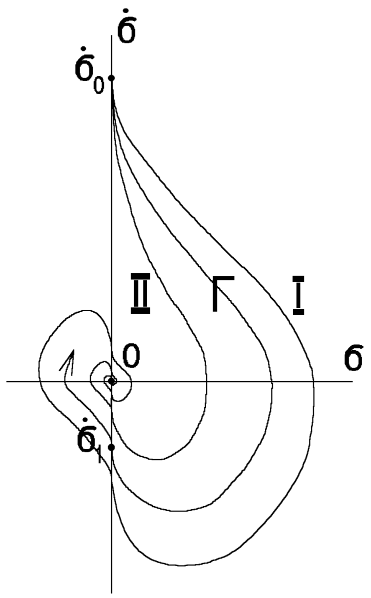


Fig. 3: Phase trajectories of the differentiator

Extending trajectory Γ, achieve sequence $\dot{\sigma}_0$, $\dot{\sigma}_1$, ... of successive intersections Γ with axis $\dot{\sigma}$. It follows from Lemma 2 below that $|\dot{\sigma}_{i+1}/\dot{\sigma}_i| <$ const < 1. Hence $\dot{\sigma}_i$ converge to 0 as a geometric progression. Also, calculating trajectory *I* approximately, achieve that (6) implies $|\dot{\sigma}_{i+1}/\dot{\sigma}_i| <$ const < 1. The last point to be checked here is the finite time convergence of the algorithm.

The state coordinates of (3), (4) are x and u_1. Consider new coordinates $\sigma = x - f(t)$ and $\xi = u_1 - \dot{f}(t)$. In these coordinates system (3), (4) takes on the form

$$\dot{\sigma} = -\lambda\,|\sigma|^{\frac{1}{2}}\,\mathrm{sign}\,\sigma + \xi,$$
$$\dot{\xi} = -\ddot{f}(t) - \alpha\,\mathrm{sign}\,\sigma,$$

and the corresponding differential inclusion is

$$\dot{\sigma} = -\lambda\,|\sigma|^{\frac{1}{2}}\,\mathrm{sign}\,\sigma + \xi, \tag{12}$$
$$\dot{\xi} \in -[\alpha - C, \alpha + C]\,\mathrm{sign}\,\sigma. \tag{13}$$

As follows from (12), $\xi = \dot{\sigma}$ when $\sigma = 0$. Hence, as follows from (13), the convergence time is estimated by the inequality

$$T \le \frac{2}{\alpha - C}\sum_{i=0}^{\infty}|\dot\sigma_i| . \qquad \square$$

Lemma 2. *The ratio of successive intersections of trajectory I with axis $\dot\sigma$ is constant and coincides with the value of function* $\Phi(\alpha,\lambda,C)$.

Proof. Let G_η, $\eta > 0$ be an operator constituted by a combination of the linear coordinate transformation g_η: $(\sigma,\xi) \mapsto (\eta^2\sigma,\eta\xi)$ and the time transformation $t \mapsto \eta t$. It transfers any vector v from the tangential space at the point $\zeta_0 = (\sigma_0,\xi_0)$ into the vector $\eta^{-1}[\frac{dg_\eta}{d\varsigma}(\zeta_0)]v$ at the point $g_\eta(\zeta_0)$. It is easily checked that (12), (13) is invariant with respect to this transformation.

Introducing new coordinates $\Sigma = \sigma/\lambda^2$ and $\Psi = \xi/\lambda^2$ and taking sign $\sigma = 1$ for positive σ, achieve that (5) describes trajectory I with some special initial conditions. Let $\xi_0=\dot\sigma_0$, $\xi_1=\tilde{\dot\sigma}_1$ be the intersections of I with the axis $\sigma = 0$. It follows from the invariance of the system with respect to transformation G_η that the value of $|\tilde{\dot\sigma}_1/\dot\sigma_0|$ does not depend on the initial value $\dot\sigma_0$. Hence, $\Phi = |\Psi(t_*)| = |\tilde{\dot\sigma}_1/\dot\sigma_0|$, $\Phi \ge |\dot\sigma_1/\dot\sigma_0|$. $\square$

Proofs of Theorems 2, 3. Consider the case when both measurement noises and discrete measurements are present. Let the input signal consist of a base signal $f_0(t)$ and a measurement noise not exceeding ε in absolute value, $\sigma = x - f_0(t)$ and $\xi = u_1 - \dot f_0(t)$. Let also the measurements be carried out at discrete times with time interval τ and $t\in[t_i, t_{i+1})$ where t_i, t_{i+1} are successive measurement times and t is the current time. Then the following differential inclusion holds:

$$\dot\sigma \in -\lambda\,|\sigma_i + [-\varepsilon, \varepsilon]|^{\frac{1}{2}}\operatorname{sign}(\sigma_i + [-\varepsilon, \varepsilon]) + \xi, \tag{14}$$

$$\dot\xi \in -[\alpha - C, \alpha + C]\operatorname{sign}(\sigma_i + [-\varepsilon, \varepsilon]),\ \ t_{i+1} - t_i = \tau. \tag{15}$$

If the continuous measurement case is considered, any appearance of indices and τ, t is to be omitted in (14), (15).

With ε, τ being zero, inclusion (14), (15) coincides with (12), (13), whose trajectories converge in finite time to the origin. It is easily seen that this implies the existence of a bounded invariant set attracting all the trajectories in finite time, when ε, τ are small. All that is needed now is to show that its size has the asymptotics defined by the statements of the Theorems when $\varepsilon, \tau \to 0$.

Applying operator G_η (see the Lemma 2 proof) to (14), (15) achieve in the coordinates $\Sigma = \sigma/\lambda^2$ and $\Psi = \xi/\lambda^2$

$$\dot{\Sigma} \in -|\Sigma_i + \tfrac{1}{\lambda^2}[-\eta^2\varepsilon, \eta^2\varepsilon]|^{\frac{1}{2}} \mathrm{sign}(\Sigma_i + \tfrac{1}{\lambda^2}[-\eta^2\varepsilon, \eta^2\varepsilon]) + \Psi, \qquad (16)$$

$$\dot{\Psi} \in -\tfrac{1}{\lambda^2}[\alpha - C, \alpha + C]\,\mathrm{sign}(\Sigma_i + \tfrac{1}{\lambda^2}[-\eta^2\varepsilon, \eta^2\varepsilon]), \quad t_{i+1} - t_i = \eta\tau. \qquad (17)$$

Consider the case when $\varepsilon = 0$ (Theorem 3). Fix some values of the expressions $\frac{\alpha - C}{\lambda^2}$ and $\frac{\alpha + C}{\lambda^2}$ and measurement step τ_0. Let the attraction set of (14), (15) be given by the inequalities $|\Sigma| \le k_1$, $|\Psi| \le k_2$. Let τ be another measurement step. After transformation G_η with $\eta = \tau/\tau_0$ achieve that $|\sigma| \le \lambda^2 k_1 \eta^2 = \lambda^2(k_1/\tau_0^2)\tau^2$, $|\xi| \le \lambda^2 k_2 \eta = \lambda^2(k_2/\tau_0)\tau$. Here k_2/τ_0, k_1/τ_0^2 are some constants. Theorem 3 follows now from the equality $\dot{\sigma} = -\lambda|\sigma_i|^{\frac{1}{2}} \mathrm{sign}\,\sigma_i + \xi$.

To prove Theorem 2 achieve from (16), (17) the following inclusion (continuous measurements are considered):

$$\dot{\Sigma} \in -|\Sigma + \tfrac{1}{\lambda^2}[-\eta^2\varepsilon, \eta^2\varepsilon]|^{\frac{1}{2}} \mathrm{sign}(\Sigma + \tfrac{1}{\lambda^2}[-\eta^2\varepsilon, \eta^2\varepsilon]) + \Psi, \qquad (18)$$

$$\dot{\Psi} \in -\tfrac{1}{\lambda^2}[\alpha - C, \alpha + C]\,\mathrm{sign}(\Sigma + \tfrac{1}{\lambda^2}[-\eta^2\varepsilon, \eta^2\varepsilon]). \qquad (19)$$

Fix some values ε_0, λ_0, α_0, C_0 and let the attraction set of (18), (19) with $\eta=1$ be given by the inequalities $|\Sigma| \le k_1$, $|\Psi| \le k_2$. Consider other values ε, λ, α, C, keeping the ends of the segment $\frac{1}{\lambda^2}[\alpha - C, \alpha + C]$ fixed. After transformation G_η with $\eta = (\varepsilon/\varepsilon_0)^{\frac{1}{2}}\lambda_0/\lambda$ achieve that $|\sigma| \le \lambda^2 k_1 \eta^2 = \lambda^2(k_1\lambda_0^2/\varepsilon_0)\varepsilon/\lambda^2 = (k_1\lambda_0^2/\varepsilon_0)\varepsilon$, $|\xi| \le \lambda^2 k_2 \eta = \lambda(k_2\lambda_0/\varepsilon_0^{\frac{1}{2}})\varepsilon^{\frac{1}{2}}$. Here $k_1\lambda_0^2/\varepsilon_0$, $k_2\lambda_0/\varepsilon_0^{\frac{1}{2}}$ are some constants. Taking into account that

$$\dot{\sigma} \in -\lambda|\sigma + [-\varepsilon, \varepsilon]|^{\frac{1}{2}} \mathrm{sign}(\sigma + [-\varepsilon, \varepsilon]) + \xi$$

achieve the statements of Theorem 2. □