The chattering problem in sliding mode systems

J. Guldner  
BMW Technik GmbH  
Hanauer Strasse 46  
D-80788 München  
Germany  
juergen.guldner@bmw.de  

V. I. Utkin  
Dept. Of Electrical Eng., The Ohio State Univ.  
205 Dreese Lab, 2015 Neil Ave.  
Columbus, OH 43210-1272  
USA  
utkin@ee.eng.ohio-state.edu

Abstract

Almost ever since sliding mode ideas have been put forward, the audible noise some sliding mode controllers exhibit has irritated control engineers and often has led to resentments and even rejection of the technique. The phenomenon is best known as “chattering”. Two main causes have been identified: First, fast dynamics in the control loop which were neglected in the system model, are often excited by the fast switching of sliding mode controllers. Second, digital implementations in micro-controllers with fixed sampling rates may lead to discretization chatter. The latter was discussed e.g. in [Utkin 1993]. In this paper, the first section, analyses the chattering phenomenon due to the first cause in detail. The subsequent sections discuss four solutions. Mathematical details can be found in [Utkin et al. 1999]

1 Introduction

The term “chattering” describes the phenomenon of finite-frequency, finite-amplitude oscillations appearing in many sliding mode implementations. These oscillations are caused by the high-frequency switching of a sliding mode controller exciting unmodeled dynamics in the closed loop. ‘Unmodeled dynamics’ may be those of sensors and actuators neglected in the principal modeling process since they are generally significantly faster than the main system dynamics. However, since ideal sliding mode systems are infinitely fast, all system dynamics should be considered in the control design.

Fortunately, preventing chattering usually does not require a detailed model of all system components. Rather, a sliding mode controller may be first designed under idealized assumptions of no unmodeled dynamics. In a second design step, possible chattering is to be prevented by one of the methods discussed Section 3 of this contribution. The solution of the chattering problem is of great importance when exploiting the benefits of a sliding mode controller in a real-life system. To some extend, chattering, without proper treatment in the control design, has been a major obstacle for implementation of sliding mode to a wide range of applications. It should be noted that the switching action itself as the core of a continuous-time sliding mode system is not referred to as chattering since in the ideal case, the switching is intended and its frequency tends to infinity; chattering, in the terminology used here, describes undesired system oscillations with finite frequency caused by system imperfections.

2 Problem Analysis

An ideal sliding mode system is depicted in Figure 1. For this example, a simulation is depicted in Figure 2.

Figure 1: Block diagram of ideal sliding mode control loop. A discontinuous controller forces the output \(x(t)\) of the plant to exactly track the desired trajectory \(x_d(t)\). No chattering occurs since the control loop is free of unmodeled dynamics.

![Block diagram of ideal sliding mode control loop.](image1.png)

Figure 2: Ideal sliding mode in first order system. State \(x(t)\) converges to desired \(x_d(t)\) in finite time, i.e. \(s(t) = 0\) after \(t = 0.45\) sec. Thereafter, control \(u(t)\) switches with infinite frequency and shows as a black area. Equivalent control \(u_{eq}(t)\), the average of discontinuous control, is drawn as a dashed line.

![Ideal sliding mode in first order system.](image2.png)
In practical applications, unmodeled dynamics in the closed loop like actuators often prevent ideal sliding mode to occur and cause fast, finite amplitude oscillations. Figure 3 shows a block diagram of the closed control loop including the previously neglected actuator dynamics. An example simulation is shown in Figure 4.

\[
\begin{align*}
\dot{x} &= w \\
\dot{w} &= v \\
\dot{v} &= -\frac{2}{\mu} v - \frac{1}{\mu^2} w + \frac{1}{\mu^2} u.
\end{align*}
\]

For the control \( u = -M \text{ sign}(x) \), the sign-varying Lyapunov function

\[
V = xv - 0.5w^2
\]

has a negative time-derivative

for small magnitudes of \( v \) and \( w \). This means that the motion is unstable in an \( \epsilon(\mu) \)-order vicinity of the manifold \( s(x) = x = 0 \).

The motion equations (1) may also be written as

\[
\dot{x}^* = -M \text{ sign}(x)
\]

\[
\mu^2 \ddot{x} + 2\mu \dot{x} + x = x^*.
\]

Sliding mode can not occur in the systems since the time derivative \( \dot{x} \) is a continuous time function and can not have its sign opposite to \( x \) in the vicinity of the point \( x = 0 \) where the control undergoes discontinuities.

To qualitatively illustrate the influence of unmodeled dynamics on the system behavior, consider a simple case with \( x_d(t) = 0 \) and motion equations

**Figure 3: Control loop with actuator dynamics neglected in ideal control design. Sliding mode does not occur since the actuator dynamics are excited by the fast switching of the discontinuous controller, leading to chattering in the loop.**

![Figure 3: Block diagram of the closed control loop including the previously neglected actuator dynamics.](image)

**Figure 4: Chattering in first order system with second order actuator dynamics under discontinuous control. After switches in control \( u(t) \), actuator output \( w(t) \) lags behind, leading to oscillatory system trajectories.**

![Figure 4: Graph of control inputs and sliding variable.](image)

**3 Chattering Suppression Methods**

The next sections discuss four solutions:

- The boundary layer solution: a continuous approximation of the discontinuity.
- The observer-based solution: generating sliding mode in a observer loop without unmodeled dynamics.
- The regular form solution: limiting sliding mode to an inner control loop of a cascaded control structure.
- The disturbance rejection solution: generating integral sliding mode in an auxiliary control loop.
3.1 Boundary layer solution

The boundary layer solution, proposed e.g. by [Slotine and Sastry 1983] and [Slotine 1984], seeks to avoid control discontinuities and switching action in the control loop. The discontinuous control law is replaced by a saturation function which approximates the sign(s) term in a boundary layer of the sliding manifold $s(t) = 0$. Numerous types of saturation functions $sat(s)$ have been proposed in the literature.

“In the large”, i.e. for $|s(t)| > \varepsilon$, $sat(s) = sign(s)$. However, in a small $\varepsilon$-vicinity of the origin, the so-called boundary layer, $sat(s) \neq sign(s)$ is continuous. As an illustrative example, consider a simple linear saturation function

$$u(t) = \begin{cases} M \text{sign}(s(t)) & \text{for } |s(t)| > \varepsilon \\ M \frac{s(t)}{\varepsilon} & \text{for } |s(t)| \leq \varepsilon \end{cases}$$

with linear proportional feedback gain $\frac{M}{\varepsilon}$ within the boundary layer in the vicinity of the origin, $|s(t)| \leq \varepsilon$, and symmetrically saturated by $M$ for $|s(t)| > \varepsilon$ outside the boundary layer. A block diagram of the example system under the above control is shown in Figure 5; simulation results are displayed in Figure 6.

![Figure 5: Saturation function replaces discontinuous controller. Instead of achieving ideal sliding mode, the system trajectories are confined to a boundary layer of the manifold $s(t) = 0$.](image)

One of the benefits of the boundary layer approach is that sliding mode control design methodologies can be exploited to derive a continuous controller. The invariance property of sliding mode control is partially preserved in the sense that the system trajectories are confined to a $\delta(\varepsilon)$-vicinity of the sliding manifold $s(t) = 0$, instead of exactly to $s(t) = 0$ as in ideal sliding mode. Within the $\delta(\varepsilon)$-vicinity, however, the system behavior is not determined, i.e. further convergence to zero is not guaranteed. This type of control design is part of a class of robust controllers which satisfy the “globally uniform ultimate boundedness” condition proposed by [Leitmann 1981]. Note that no real sliding mode takes place since the switching action is replaced by a continuous approximation.

![Figure 6: Saturation function approximating control discontinuity in boundary layer yields chattering free system trajectories. State $x(t)$ converges to desired $x_d(t)$, but does not track exactly as in ideal sliding mode.](image)

3.2 Observer-based solution

The boundary layer approach avoids generating sliding mode by replacing the discontinuous switching action by a continuous saturation function. In many applications, however, control discontinuities are inherent to the system, e.g. in various voltage inputs of power converters or electric drives. When implementing a continuous controller, a technique like pulse-width modulation (PWM) has to adapt the control law to the discontinuous system inputs. In the light of recent advances of high-speed circuitry, it seems unjustified to by-pass a system’s discontinuous control inputs by converting a continuous controller e.g. via a PWM scheme. Rather, such system specifications call for alternative methods to prevent chattering while preserving control discontinuities.

An asymptotic observer in the control loop can eliminate chattering despite discontinuous control laws. The key idea as proposed by [Bondarev et al. 1985] is to generate ideal sliding mode in an auxiliary observer loop rather than in the main control loop. Ideal sliding mode is possible in the observer loop since it is entirely generated in the control software and thus does not contain any unmodeled dynamics. The main loop follows the observer loop according to the observer dynamics. Despite applying a discontinuous control signal with switching action to the plant, no chattering occurs and the system behaves as if an equivalent continuous $u_{eq}(t)$ control was applied. The definition of equivalent control can be found e.g. in [Utkin et al. 1999]. Figure 7 shows a block diagram, simulation results can be found in Figure 8.
Figure 7: Control loop with auxiliary observer loop. Ideal sliding mode occurs in observer manifold $\dot{x}_\text{st} = 0$ since the observer loop is free of unmodeled dynamics.

Figure 8: Observer in auxiliary control loop enables chattering free system trajectories despite discontinuous control switching after sliding manifold $\dot{x}_\text{st} = 0$ is reached. The plant output $x(t)$ follows the observer output $\hat{x}(t)$ without chattering despite discontinuous control $u(t)$ applied to main loop with actuator dynamics.

3.3 Regular form solution

Both the boundary layer approach and the observer-based solution to the chattering problem assume that the ‘unmodeled’ dynamics are completely unknown. In practical applications, however, at least partial information about unmodeled dynamics, in particular of actuators, is often available together with measurements of the actuator outputs. For example, for electric drives, models are readily available, but may contain uncertain parameters. Thus in the design of a controller for the overall system, these dynamics can be included into the control design to enhance the performance of the overall system.

Since the actuator dynamics and the plant dynamics are block separated, i.e. the output(s) of the actuator(s) are the input(s) of the plant, a cascaded control structure can be designed following the regular form approach or the block control principle (see also [Drakunov et al. 1990]). The basic idea is to design a cascaded controller (see Figure 9) in two steps. In the first step, a continuous controller is derived for the plant under the assumption that the plant input(s) is/are the actual control input(s) to the overall system, defining ‘desired’ actuator output(s) $w_d(t)$. In the second step, the actuator input(s) $u(t)$, i.e. the real control input(s) of the system, is/are used to ensure the actuator output(s) track the desired output(s) exactly via sliding mode control with $w(t) = w_d(t)$. This approach is a special case of cascaded control structures as applied the block control principle, see e.g. [Drakunov et al. 1990] and the integrator backstepping method, see e.g. [Krstic et al. 1995]. See Figure 10 for simulation results.

Figure 9: Cascaded controller with continuous auxiliary control and discontinuous actuator control loop.

Figure 10: Linear feedback controller with feedforward of desired trajectory, $u_c(t)$, leads to inaccurate tracking of desired trajectory $x_d(t)$, since the closed loop system (8.5.3) is perturbed by unknown plant dynamics and external disturbance $d(x,t)$. 
3.4 Disturbance rejection solution

The regular form solution in the previous section relies on a continuous controller to achieve tracking of the desired trajectory \( x_d(t) \) by the output \( x(t) \) of the plant. The linear controller usually is augmented by an estimate of the disturbance. Often, such an estimate is not readily accessible. The disturbance rejection approach discussed in this section provides means to obtain an accurate disturbance estimate while avoiding chattering in the main control loop. This approach can be viewed as a special case of so-called integral sliding mode. A more mathematical background of integral sliding mode is described e.g. in [Utkin and Shi 1996, Utkin et al. 1999].

The main idea of disturbance rejection via sliding mode is to compose the overall controller of a continuous part and of a discontinuous part. The continuous component is used to control the overall behavior of the system while the discontinuous component is used to reject disturbances and to suppress parametric uncertainties. A block diagram is shown in Figure 11, with simulation results being displayed in Figure 12.

Figure 11: Disturbance rejection via sliding mode with auxiliary controller loop to avoid chattering. A continuous controller \( u_c(t) \) is augmented by a disturbance rejection controller \( u_d(t) \), derived from a low pass filtered discontinuous controller for an auxiliary control variable \( z(t) \).

Figure 12: Performance of linear feedback controller with feed-forward of desired trajectory, \( u_c(t) \), is significantly improved by disturbance rejection controller \( u_d(t) \) based on auxiliary variable \( z(t) \) for estimating unknown plant dynamics and external disturbances, summarized as \( f(x,t) \).

4 Example: Automatic Steering Control

Automation of vehicles, e.g. for Automated Highways Systems (AHS), has been discussed for several decades and is studied in various programs worldwide in the framework of Intelligent Transport Systems, see e.g. [Stevens 1996] or [Tsugawa 1996].

Two control subtasks arise for automated driving: Steering control to keep the vehicle in the lane (controlling the lateral motion) and throttle/brake control to maintain speed and proper spacing between vehicles (controlling the longitudinal motion). Both subtasks have been solved using sliding mode control. The focus in this section will be on automatic steering control. Longitudinal control was studied e.g. by [Hedrick et al. 1994] or [Pham et al. 1994].

The automatic steering system of an automated vehicle consists of a reference system to determine the lateral vehicle position with respect to the lane center, sensors to detect the vehicle motion (typically yaw rate and lateral acceleration), and a steering actuator to steer the front wheels. The variety of employed reference systems ranges from look-ahead systems like machine vision or radar to look-down systems like electric wires or magnets embedded in the road surface. ‘Look-ahead/look-down’ describes the point of measurement of lateral vehicle displacement from the reference to be ahead of the vehicle or directly down from the front bumper (see [Patwardhan et al. 1997] for a more detailed treatment). The control design below, however, is valid for any reference system.

Control design is usually based on the so-called single track model, which concentrates on the main vehicle mass by lumping the two wheels at each axle into a single wheel.
The road/tire interaction forces are responsible for generating planar lateral and yaw vehicle motions, with the front wheel steering angle \( \delta_f \) being the input variable. A linearized second order model for constant speed \( v \) is given by

\[
\begin{pmatrix}
\dot{\beta} \\
\dot{\psi}
\end{pmatrix} = 
\mu
\begin{bmatrix}
\frac{-c_f + c_s}{Mv} & -1 - \frac{c_f l_f - c_s l_s}{Mv^2} \\
\frac{c_f l_f - c_s l_s}{J} & -\frac{c_f l_f^2 + c_s l_s^2}{Jv}
\end{bmatrix}
\begin{pmatrix}
\beta \\
\psi
\end{pmatrix} + 
\mu
\begin{bmatrix}
\frac{c_f}{Mv} \\
\frac{c_f l_f}{J}
\end{bmatrix}
\delta_f
\tag{5}
\]

with states side slip angle \( \beta \) and yaw rate \( \psi \). For a detailed derivation of (5) see e.g. [Peng 1992]. Parameters are vehicle mass \( M \) and yaw inertia \( J \), distances \( l_f \) and \( l_r \) of front and rear tires, and road adhesion factor \( \mu \). All parameters are uncertain within known bounds, e.g. \( 0 < \mu^- \leq \mu \leq \mu^+ \leq 1 \).

Figure 13: Single track model of a vehicle following a lane reference. Sensors at the front and tail bumpers measure lateral displacements \( y_{es} \) and \( y_{eS} \), respectively. Also shown are vehicle states, side slip angle \( \beta \) and yaw rate \( \psi \), input steering angle \( \delta_f \), and various distances from center of gravity CG.

When following a reference path with curvature \( \rho_{ref} \), as depicted in Figure 13, lateral vehicle displacement \( y_{es} \), measured at some sensor position \( d_x \) ahead of CG, and angular error \( \psi \) can be described by linearized dynamic model

\[
\begin{align*}
\dot{y}_{es} &= v(\beta + \psi_e) + d_x \psi, \\
\dot{\psi}_e &= \psi - v \rho_{ref}.
\end{align*}
\tag{6}
\]

Given (5) and (6), various control design options are possible. As an example, we present a cascaded control design under the assumption that vehicle yaw rate \( \psi \) is measurable by a gyroscope. The control design follows the regular form methodology (see Section 3.3) and considers subsystem (5) as the input to subsystem (6). Hence, the first design step assumes yaw rate \( \psi \) to be a direct input to (6) and derives a desired yaw rate \( \psi_d \). The second step then ensures that the actual, measured vehicle yaw rate \( \psi \) follows \( \psi_d \) exactly via appropriate control design for steering angle \( \delta_f \) in (5), the true system input. A suitable continuous feedback/feedforward “yaw rate” controller to stabilize the first equation in (6) would be

\[
\dot{\psi}_d = -\frac{1}{I_s} \left( v(\beta + \psi_e) + Cy_{es} \right),
\tag{7}
\]

with linear feedback gain \( C > 0 \). However, neither side slip angle \( \beta \) nor yaw angle error \( \psi_e \) can be measured and hence have to be estimated by an observer (for details, see e.g. [Guldner et al. 1994, Ackermann et al. 1995]). Introducing auxiliary variable \( \hat{z} = \hat{\beta} + \psi_e \), an observer is designed as

\[
\begin{bmatrix}
\dot{\hat{y}}_{es} \\
\dot{\hat{z}}
\end{bmatrix} =
\begin{bmatrix}
0 & v \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\hat{y}_{es} \\
\hat{z}
\end{bmatrix} +
\begin{bmatrix}
I_s \\
0
\end{bmatrix}
\begin{bmatrix}
\hat{c}_1 \\
\hat{c}_2
\end{bmatrix}
\begin{bmatrix}
\hat{y}_{es} \\
\hat{z}
\end{bmatrix},
\tag{8}
\]

with feedback of the observation error \( \bar{y}_{es} = y_{es} - \hat{y}_{es} \) via gains \( \hat{c}_1 > 0 \) and \( \hat{c}_2 > 0 \), chosen faster than the vehicle dynamics in (5). With the help of the observed auxiliary variable \( \hat{z} = \hat{\beta} + \psi_e \), a desired yaw rate is defined as

\[
\psi_d = -\frac{1}{I_s} \left( v(\hat{\beta} + \psi_e) + Cy_{es} \right).
\tag{9}
\]

The second step of control design uses the steering angle \( \hat{\delta}_f \) as the input to (5) to drive yaw rate error \( \psi_e = \psi_d - \psi \) to zero, e.g. by purely discontinuous sliding mode control

\[
\hat{\delta}_f = \delta_0 \text{sign} \psi_e.
\tag{10}
\]

The stability analysis follows the previously discussed Lyapunov approach and is omitted here for brevity. Alternatively to (10), a combination of continuous feedback/feedforward and a discontinuity term could be employed, i.e.

\[
\delta_f = \delta_0 \text{sign} \psi_e + \frac{1}{\hat{c}_1 I_f} \left( \left( \hat{c}_1 \hat{\beta} - \hat{c}_1 \hat{\beta} \right) + \frac{1}{\hat{c}_1} \left( \hat{c}_1 \hat{\beta} \right) \frac{1}{\hat{\beta}} \psi_d \right),
\tag{11}
\]

where estimates of vehicle parameters are denoted with ‘hats’ (\( \hat{\cdot} \)), the estimate of side slip angle, \( \hat{\beta} \), stems from an observer similar to (8) and the derivative of the desired yaw rate, \( \hat{\psi}_d \), can be derived from (9) by virtue of known observer dynamics. Due to the continuous feedback/feedforward terms in (11), the gain of the discontinuity term can be reduced as compared to (10), i.e. \( \delta_1 < \delta_0 \).

The above control design neglects the dynamics of the steering actuator, which will lead to chattering in practical implementations. In addition to the chattering prevention methods discussed above, the introduction of an integrator in the control loop proved to be a promising approach. Originally, the integrator was a physical model of the actuator dynamics (see e.g. [Ackermann et al. 1993]) with the steering rate \( u \) being the system input as
Define a second order sliding variable
\[ s = C_3 \dot{\psi}_f + \ddot{\psi}_f, \]  
(13)
leading to a control law
\[ \delta_f = \delta_0 \text{sign} s \]  
(14)
instead of (10). The alternative feedback/feedforward controller term in (11) has to be adjusted accordingly. If the real steering actuator is not an integrator as in (12), but features more complex dynamics e.g. of an electro-hydraulic actuator, two design alternatives are left to the control engineer. Either, a sliding mode controller is designed according to (10/11) with appropriate measures to prevent chattering as discussed above, or an integrator like (12) is introduced as part of the controller, i.e. realized in the controller software. The latter case follows the ideas of integral sliding mode by implementing sliding motion in an integral manifold rather than directly in the control input variable \( \delta_f \). Hence the switching action of the sliding mode discontinuity is first filtered by integrator (12) and thus does not directly reach the input \( \delta_f \), which inherently prevents chattering. A different integrator location in the control loop was proposed by [Pham et al. 1994], before rather than after the switching discontinuity. The interested reader is referred to [Hingwe and Tomizuka 1995] for a comparison of different integrator locations in the controller loop. Experimental results from this work are displayed in Figure 14.

\[ \dot{\delta}_f = u, \]  
(12)
rather than the steering angle \( \delta_f \) itself. The additional integrator only requires to alter the outer control loop (10/11).

5 Discussion

In applications of sliding mode control, unmodeled dynamics in the control loop are often excited by the discontinuous switching action of a sliding mode controller, leading to oscillations in the motion trajectory. Due to the acoustic noise such oscillations may cause in mechanical systems, this phenomenon is also referred to as ‘chattering’. This paper studied the chattering problem and presented four solutions. All four discussed solutions to the chattering problem reliably eliminate chattering in the control loop. In order to successfully prevent chattering, all methods require some estimate of the time constant or the bandwidth of the unmodeled dynamics. Instead of achieving exact tracking performance as in ideal sliding mode, small tracking error are tolerated.

In general, the achievable performance of a control system depends on the performance of sensors and actuators, availability of knowledge about the system, i.e. the quality of the system model, and the availability of measurements of system variables. For example, a system with a slow actuator can not fully reject fast disturbances, regardless of the control design methodology used. A sliding mode controller under ideal conditions is able to fully exploit the system capabilities. Under realistic conditions, a chattering prevention scheme should be selected depending on the system specifications to ensure good system performance.

The first of the discussed method substitutes the discontinuity of a sliding mode controller by a saturation function and yields motion in a boundary layer of the sliding manifold instead of true sliding along the manifold. Effectively, sliding mode methodology is utilized to design a continuous high-gain controller which respects bounds on the control resources.

The second method shifts the switching action of the sliding mode controller into an auxiliary observer loop, thus circumventing unmodeled dynamics in the main loop and achieving ideal sliding mode in the observer loop. The plant follows the ideal trajectory of the observer according to the observer performance. Since the control input to the plant is still discontinuous, this method is ideal for systems which already have an observer in the control structure or for systems with inherently discontinuous control inputs like voltage inputs of electric drives. Implementation of a continuous controller in a system with discontinuous inputs generally requires pulse-width modulation (PWM), whereas direct implementation of sliding mode control with an observer avoids the detour via PWM.

The third method is mainly designed for systems where some knowledge of the unmodeled dynamic and intermediate measurements are available, e.g. known actuator dynamics. Such systems consisting of separated blocks may be controlled with a cascaded control structure which avoids chattering by explicitly taking the unmodeled dynamics into account for the control design. In this sense, they are no longer “unmodeled”, but rather part of the overall system model.

\[ \text{Figure 14: Experimental results of an automatic steering controller based on sliding mode design.} \]  
\footnote{1 Courtesy Dr. Hingwe and Prof. Tomizuka.}
The last method combines a continuous and a discontinuous controller to achieve good performance without chattering. The continuous part controls the overall motion whereas the task of the discontinuous part is to reject the influence of parametric uncertainty and disturbances. This method is a special case of integral sliding mode and is especially useful for systems with large uncertainties and/or disturbances.

All four methods possess their advantages and disadvantages which depend on the system specifications. When designing a sliding mode controller for a given system, the choice of which method to employ to prevent chattering usually requires careful consideration of all details; unfortunately, no textbook solution exists to cope with all system in a general manner.

References


Peng, H., Vehicle Lateral Control for Highway Automation, Ph.D. Dissertation, University of California (Dept. of Mechanical Engineering), Berkeley, CA, USA, 1992


Stevens, W., “The Automated Highway Systems Program: A Progress Report”, in Preprints of the 13th IFAC World Congress (plenary volume), San Francisco, CA, USA, pp. 25-34, 1996.


