Sliding modes in intersecting switching surfaces

James C. Alexander
Department of Mathematics
University of Maryland College Park
College Park, MD 20742-4015, USA
e-mail: ⟨jca@math.umd.edu⟩

Thomas I. Seidman
Department of Mathematics and Statistics
University of Maryland Baltimore County
Baltimore, MD 21228, USA
e-mail: ⟨seidman@math.umbc.edu⟩

September 14, 1995

ABSTRACT: When a flow, discontinuous across a switching surface, points 'inward' so one cannot leave, it induces a unique flow within the surface, called the *sliding mode*. When several such surfaces intersect, one would seek a flow within the intersection, but some difficulties arise. We explore here the extent of the ambiguity involved in this situation and then show that for a certain form of 'natural mechanism for implementation' (sigmoid blending) one does indeed inherit, as a residual effect of this implementation, sufficient information to characterize a well-defined sliding mode in the intersection of two switching surfaces.

KEY WORDS: *sliding mode, differential equation, hybrid system, asymptotics.*

1. Introduction

In [3], Filippov, from a dynamical-systems point of view, and in [7], Utkin, from a control-theory point of view, considered certain hybrid differential systems with particular discontinuities in the right-hand side. Such discontinuities occur across smooth hypersurfaces, and they model systems which change sharply across such hypersurfaces. If the dynamics leads the evolution of the system into one or more of these hypersurfaces, the determination of a *sliding mode*[1] of the system is an issue to be addressed. In [3], Filippov derived a general condition such sliding modes must satisfy, and this condition is sufficient to specify the mode if there is only one discontinuity hypersurface. If there are more than one, his condition must be supplemented — by what, depends on the particular mechanism being modelled.

We are concerned with local determination of the sliding mode for a hybrid system

$$\dot{x} = f_{j(x)}(x) \tag{1.1}$$

on a finite- or sometimes infinite-dimensional state space $\mathcal{X}$ where each mode f_j is smooth and the open domains $\Omega_j = \{\xi : j(\xi) = j\}$ are separated by unions of smooth switching surfaces; note that initially there is no specification for the vector field giving $\dot{x}$ (as a so-called 'sliding mode') when x actually lies *on* one of these surfaces of discontinuity. We assume that (locally) the switching surfaces are codimension 1 manifolds $\mathcal{Z}_k$ and that we are concerned with appropriate determination of the sliding mode in an intersec-

[1] Note that (1.1) does not assert a specification of the field *on* a surface of discontinuity. When a trajectory crosses, the value at such an isolated point is irrelevant, but if the vector fields on either side point 'inward' — so an arriving trajectory cannot escape — then one would be constrained to follow some flow within the surface, called the '*sliding mode*, and the specification of this flow is our principal concern here.

Such modal systems often arise in contexts of control, for which the mode selection $j = j(x)$ represents the intentional imposition of a control policy. We note, in particular, that this control strategy may be specifically intended [7] to take advantage of the properties of the resulting sliding mode. Thus, the situation we consider represents an extension of the existing theory of [3], [7], etc., which may lead to useful new control strategies; cf., e.g., [4]. Now, in comparison with [3], [7], some of the details of introducing a specific physical mechanism to implement the switching implicit in the selection $j = j(x)$ remain significant in selecting the sliding mode within the intersection of switching surfaces.

tion manifold

$$\mathcal{Z} := \bigcap_{k=1}^{m} \mathcal{Z}_k. \tag{1.2}$$

Thus, we ask: *"How should we* define $f(\xi_*)$ *for a given* $\xi_* \in \mathcal{Z}$*?"*

In many contexts it is reasonable to interpret (1.1) through the weak formulation

$$x(t) = \xi_0 + \int_0^t \varphi(s)\,ds \qquad \text{with } \varphi(s) \in \mathcal{H}(x(s)) \text{ ae} \tag{1.3}$$

where $\mathcal{H}(\xi)$ is just the hull (convex combinations) of the adjacently available original field values of (1.1).

This use of $\mathcal{H}$ does not cover some physically interesting situations as, e.g., dry friction. Nevertheless — and this viewpoint dominates our discussion — the formulation (1.3), correctly describes the results of limit processes corresponding to a variety of physical implementations of (1.1). In particular, the use of $\mathcal{H}$ here is a consequence of the 'localization principle'

$$\begin{array}{c}\textit{If } f(\eta) \in \hat{\mathcal{C}} \textit{ for some convex set } \hat{\mathcal{C}} \subset \mathcal{X} \textit{ and all } \eta \textit{ in some neighborhood of } \xi, \\ \textit{then } \mathcal{F}(\xi) \subset \hat{\mathcal{C}}.\end{array} \tag{1.4}$$

We emphasize that our concern is with uniqueness. The *existence* of weak solutions for (1.3) is not at issue: one may define a 'Filippov ε-solution' of (1.1) as a function $x_\varepsilon(\cdot)$ satisfying

$$x_\varepsilon(t) = \xi_0 + \int_0^t \varphi(s)\,ds \qquad \text{with } \varphi(s) \in \mathcal{H}_\varepsilon(x_\varepsilon(s)) \text{ ae} \tag{1.5}$$

where

$$\mathcal{H}_\varepsilon(\xi) := \{\varphi \in \mathcal{X} : \exists j,\ \eta \in \Omega_j \text{ such that } |\eta - \xi| \le \varepsilon,\ |f_j(\eta) - \varphi| \le \varepsilon\}; \tag{1.6}$$

note that $\mathcal{H}(\xi) = \bigcap_{\varepsilon > 0} \overline{hull}\mathcal{H}_\varepsilon(\xi)$. Under very mild hypotheses one can show — even in the infinite dimensional case; cf., [5] — that there always exist convergent sequences of such ε-solutions with $\varepsilon = \varepsilon_n \to 0$ and that the limit of such ε-solutions is a solution of (1.3); compare Theorem 2 below.

We are concerned with a situation in which any trajectory through ξ_* must lie locally in the manifold $\mathcal{Z}$. In this case $\dot{x}$ must be ae in the tangent space $\mathcal{T}(\xi_*)$ — cf., (2.10) — and we are led to define the *set of Filippov vectors* at $\xi \in \mathcal{Z}$ as

$$\mathcal{F}(\xi) := \mathcal{H}(\xi) \cap \mathcal{T}(\xi). \tag{1.7}$$

We refer to the replacement of $\mathcal{H}$ by $\mathcal{F}$ in (1.3) as the 'Filippov condition'. For a single switching surface ($m = 1$), Filippov has observed [3] that in relevant cases this $\mathcal{F}(\xi_*)$ is a singleton so one has a unique $f(\xi_*)$ — i.e., a well-determined sliding mode and so, in general, unique solutions of (1.1). As we shall see, the uniqueness may fail for $m > 1$ and our objectives in this paper are to understand this ambiguity and to attempt to resolve it. The Filippov condition $\dot{x} \in \mathcal{F}(x)$ must then be supplemented by some selection, and there does not appear to be available any such *universal* principle as the 'localization principle' (1.4).

However, in most models, there is more structure to be found if one seeks it. Our underlying viewpoint is that a problem such as (1.1) can occur physically only if one has some specific *mechanism* for implementing the mode-switching so the formulation (1.1) is really an idealized 'reduced model' for a more complicated situation. The selection of a specific sliding mode to resolve the ambiguity may reflect the limiting procedure implicit in this idealization — indeed, the very fact of such ambiguity shows that the outcome *must* reflect this — so some modelling of the sensors/actuators which govern implementation of the control mechanisms may provide extra structure as a residual effect. Consideration of this extra structure offers the real possibility of recovering a unique recipe for a sliding mode for several switching surfaces. Working locally, i.e., in a neighborhood of some point ξ_* in the intersection of the switching surfaces, we assume that the dynamics of the controlled variables (normal to the switching surfaces) would be much faster than that along the switching surfaces (infinitely fast in the ideal limit) and we wish to utilize an asymptotic analysis of the fast dynamics.

For the intersection of two switching surfaces we are able to complete the program with respect to one particular class of implementation mechanisms. The mechanisms we consider here may be described as *sigmoid blending.* They correspond precisely to the use of sigmoid functions in the treatment of (artificial) *neural nets* to approximate dynamically (in a way which may be viewed as one implementation of 'fuzzy logic') the discontinuous ON/OFF switching of a logical circuit. Thus, one implication of our treatment addresses the asymptotics of such a neural net as these sigmoidal blending functions are taken to be more sharply ON/OFF, leading to the idealized McCulloch-Pitts neuron.

The paper is organized as follows. In Section 2, we set up the necessary background and notation. In Section 3, we discuss the ambiguity inherent in

Filippov's condition. In Section 4, we continue the discussion of the consequences of Filippov's condition, and discuss several possible mechanisms to further resolve the ambiguity. In Section 5, we turn to a specific mechanism we call blending. In this section, we prove blending gives rise to a unique well-defined sliding mode for two discontinuity hypersurfaces.

2. Background and Notation

We assume that a finite number of smooth *sensor functionals* $Y_k : \mathcal{X} \to \mathbb{R}$ would be defined, with discontinuity permitted for f only when some Y_k changes sign; thus the corresponding switching surface $\mathcal{Z}_k$ is given as the level surface

$$\mathcal{Z}_k = \{x : Y_k(x) = 0\}. \tag{2.1}$$

Our considerations are local so we work in a neighborhood $\mathcal{O}$ of some specific point $\xi_* \in \mathcal{X}$ and set

$$\mathcal{K} = \mathcal{K}(\xi) := \{k : Y_k(\xi) = 0\}. \tag{2.2}$$

Without loss of generality,[2] we may assume that for $\xi \in \mathcal{O}$ one has $\mathcal{K}(\xi) \subset \mathcal{K}(\xi_*) =: \mathcal{K}_* = \{1, \ldots, m\}$. We impose here the *nondegeneracy condition* that the normals $\{\mathbf{n}_k = \nabla Y_k(\xi)\}$ satisfy

$$\{\mathbf{n}_k(\xi) : k \in \mathcal{K}(\xi)\} \text{ is linearly independent} \tag{2.3}$$

for $\xi = \xi_*$, and so, by continuity, in $\mathcal{O}$. Hence the switching surfaces $\mathcal{Z}_k$ are locally manifolds intersecting transversely. This is only possible if $dim\,\mathcal{X} \geq m$ and is only of interest for us if $dim\,\mathcal{X} > m$: in this case the intersection $\mathcal{Z} := \{\xi : \mathcal{K}(\xi) = \mathcal{K}_*\}$ — compare (1.2) — is a manifold of codimension m in $\mathcal{X}$.

We note at this point that the nondegeneracy assumption (2.3) together with the assumed smoothness of the sensor functionals Y_k ensure that (locally) one can change variables: $x \mapsto w$ diffeomorphically so that the sensor values are simply the first m coordinates (i.e.,the switching surfaces become the first m coordinate planes in the new variables) and the complementary

[2] I.e., $\mathcal{O}$ is taken to avoid any other locally irrelevant switching surfaces and we suitably renumber the sensor functionals.

coordinate z is a local coordinatization of $\mathcal{Z}$ (equivalently, of the tangent space $\mathcal{T}$). Thus we obtain

$$(2.4)\quad x \mapsto w = W(x) := [y, z], \qquad y := (y_1, \ldots, y_m) \text{ with } y_k := Y_k(x).$$

as the full set of new coordinates; a neighborhood of ξ_* (still denoted by $\mathcal{O}$) is coordinatized by $\mathbb{R}^m \times \mathcal{T}$. Note that we then have a corresponding transformation for the vector fields: we must set

$$f^{\boldsymbol{\sigma}} \mapsto \hat{f}^{\boldsymbol{\sigma}} := \left[\frac{dW}{dx}\right] f^{\boldsymbol{\sigma}} =: [u^{\boldsymbol{\sigma}}, v^{\boldsymbol{\sigma}}] \tag{2.5}$$

(written in terms of w) in order to convert $\dot{x} = f$ to $\dot{w} = \hat{f}$. Although we do not shift to this viewpoint until later, we do note that the entire analysis of the sliding modes could be carried through in this simplified setting and then re-interpreted.

For $k = 1, \ldots, m$, let σ_k be a plus ($+$) or a minus ($-$) sign (equivalently, ± 1). Intuitively, each σ_k represents the $\pm$ state of the k^{th} sensor — e.g., denoting whether the temperature is above or is below the desired value, etc. — or the (Boolean) truth value of the logical proposition, 'the k^{th} sensor value is positive'. Let $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_m)$ be a sequence of m such signs; there are 2^m possible values of the Boolean vector $\boldsymbol{\sigma}$. Let

$$\mathcal{O}^{\boldsymbol{\sigma}} = \bigcap_{k=1}^{m} \{x \in \mathcal{O} : \operatorname{sgn} Y_k(x) = \sigma_k\} \tag{2.6}$$

so each $\mathcal{O}^{\boldsymbol{\sigma}}$ is the pre-image of an open orthant in the *sensor space* $\mathbb{R}^m$; we then set

$$f^{\boldsymbol{\sigma}} = f\big|_{\mathcal{O}^{\boldsymbol{\sigma}}} \tag{2.7}$$

and assume that each $f^{\boldsymbol{\sigma}}$ extends smoothly to $\overline{\mathcal{O}^{\boldsymbol{\sigma}}}$. Then for $\xi \in \mathcal{Z}$ one has $\varphi \in \mathcal{H}(\xi)$ if and only if there are nonnegative coefficients $\gamma_{\boldsymbol{\sigma}}$ such that

$$\varphi = \sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}} f^{\boldsymbol{\sigma}} \qquad \text{with } \sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}} = 1. \tag{2.8}$$

The situation of interest for us here is that in which the vector fields $f^{\boldsymbol{\sigma}}$ adjacent to $\xi_* \in \mathcal{Z}$ all 'point inward', i.e. the feedback of each sensor is negative. This is our key hypothesis: at ξ_* (and so locally, in $\mathcal{O}$, by continuity) one has

$$\text{sgn } (\lambda_{k,\boldsymbol{\sigma}}) = -\sigma_k \quad \text{where } \lambda_{k,\boldsymbol{\sigma}} := \mathbf{n}_k \cdot f^{\boldsymbol{\sigma}}. \tag{2.9}$$

In this case we see that a trajectory entering $\mathcal{Z}_k$ cannot leave it, but must remain in this switching surface. To satisfy (1.3) for an interval of t with x remaining in $\mathcal{Z}_k$, necessarily $\dot{x}$ is tangential to $\mathcal{Z}_k$. For this to hold for each such k, we see that $\varphi(s)$ in (1.3) is not only in $\mathcal{H}(\xi)$ but also must ae be in the tangent space: in view of (2.3),

$$\varphi(s) \in \mathcal{T}(\xi) := \{\varphi : \mathbf{n}_k \cdot \varphi = 0 \text{ for } k \in \mathcal{K}(\xi)\} \tag{2.10}$$

with $\xi = x(s)$ and $\mathbf{n}_k = \mathbf{n}_k(\xi)$ — i.e., $\varphi \in \mathcal{F}$ as in (1.7). Note that the inwardness condition (2.9) as well as the condition (2.10) depend only on the transverse components of the fields; indeed, (2.10) simply asserts that this component must vanish for a sliding mode so $\varphi^\perp = 0$ and $\varphi = \varphi^{\parallel}$.

Combining (2.8) with (2.10), we see more explicitly that for $x(s) = \xi \in \mathcal{Z}$ we should have $\varphi(s) = \sum \gamma_{\boldsymbol{\sigma}} f^{\boldsymbol{\sigma}}(\xi)$ with nonnegative coefficients $\gamma_{\boldsymbol{\sigma}}$ satisfying

$$\sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}} = 1 \quad \text{and} \qquad \sum_{\boldsymbol{\sigma}} \lambda_{k,\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}} = 0, \quad (k = 1 \ldots, m) \tag{2.11}$$

with $\lambda_{k,\boldsymbol{\sigma}} := \mathbf{n}_k \cdot f^{\boldsymbol{\sigma}}$ as in (2.9). Note that if we write $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}(\xi)$ for the $m \times 2^m$ matrix $(\lambda_{k,\boldsymbol{\sigma}})$, then (2.11) just asserts that

$$\mathbf{A}(\gamma_{\boldsymbol{\sigma}}) = \begin{pmatrix} 1 \\ 0_m \end{pmatrix} \qquad \text{with } \mathbf{A} = \mathbf{A}(\xi) := \begin{pmatrix} 1_m \\ \boldsymbol{\Lambda} \end{pmatrix}. \tag{2.12}$$

[Here 0_m is a column m-tuple of '0's and 1_m a row m-tuple of '1's.] This is sufficient to characterize the sliding mode when $m = 1$ but is only a necessary condition when $m > 1$. We define the coefficient sets $\mathcal{C} = \mathcal{C}(\xi)$ and $\mathcal{C}_+ = \mathcal{C}_+(\xi)$ by

$$\begin{aligned} \mathcal{C} \quad &:= \{(\gamma_{\boldsymbol{\sigma}}) \text{ satisfying (2.11)}\}, \subset [0,1]^{2^m} \\ \mathcal{C}_+ \quad &:= \{(\gamma_{\boldsymbol{\sigma}}) \in \mathcal{C} : \text{ each } \gamma_{\boldsymbol{\sigma}} \geq 0\} \end{aligned} \tag{2.13}$$

so $\mathcal{C}_+$ is the set of 'Filippov coefficients' and the set $\mathcal{F} = \mathcal{F}(\xi)$ of 'Filippov vectors', previously defined by (1.7), can now be represented as

$$\begin{aligned} \mathcal{F} = \mathcal{F}(\xi) \quad &:= \mathbf{F}(\mathcal{C}_+(\xi)) \\ \text{with} \qquad &\mathbf{F} = \mathbf{F}(\xi) : \mathbb{R}^{2^m} \to \mathbb{R}^m : (\gamma_{\boldsymbol{\sigma}}) \mapsto \textstyle\sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}} f^{\boldsymbol{\sigma}}, \end{aligned} \tag{2.14}$$

noting that the range of $\mathbf{F}(\xi)$ lies in the tangent space $\mathcal{T}(\xi)$. One easily sees from (2.12) and (2.14) the obvious inequalities

$$(2.15)\qquad dim\,\mathcal{F}\quad \begin{cases} \leq dim\,\left[hull_{\boldsymbol{\sigma}}\left\{(f^{\boldsymbol{\sigma}})^{\|}\right\}\right] \leq dim\,\mathcal{Z}, \\ \leq dim\mathcal{C} = dim\,\mathcal{N}(\mathbf{A}) := \nu \end{cases}$$

and, in particular, one sees that

$$(2.16)\qquad \mathcal{F}(\xi) \text{ is a singleton} \quad \Longleftrightarrow \quad ker(\mathbf{A}) \subset ker(\mathbf{F}).$$

3. Ambiguity

In this section we consider first the extent of the 'coefficient ambiguity' corresponding to the defining relations (2.11) together with positivity and then discuss the implication of this for the 'field ambiguity' of φ as an element of $\mathcal{F} = \mathbf{F}(\mathcal{C}_+)$.

THEOREM 1: *Under the inwardness assumption* (2.9), *the set* $\mathcal{C}_+(\xi)$ *of 'Filippov coefficients' is a bounded convex set with nonempty* ν-*dimensional interior in* $(0,1)^{2^m}$ *with* $\nu = 2^m - m - 1$.

PROOF: We first note that the system has at least one positive solution — i.e., that $\mathcal{C}_+$ is nonempty and, indeed, has nonempty 'interior' — by the construction of a particular $(\gamma_{\boldsymbol{\sigma}}) \in \mathcal{C}$ with each $\gamma_{\boldsymbol{\sigma}} > 0$.

Suppose we already knew how, recursively, to get positive solutions of the defining system (2.11) — for some $\hat{m}$ and all $f^{\boldsymbol{\sigma}'}$ satisfying (2.9), corresponding to each $\boldsymbol{\sigma}' \in \{\pm\}^{\hat{m}}$. For $\hat{m}+1$ we now write $\boldsymbol{\sigma} = [\boldsymbol{\sigma}', \pm]$ (and the modes $f^{\boldsymbol{\sigma}}$ as $f^{[\boldsymbol{\sigma}',\pm]}$) with $\boldsymbol{\sigma}'$ corresponding to $\hat{m}$ and $\pm = \sigma_{\hat{m}+1}$. By the inductive hypothesis, we can obtain (separately) coefficients $\gamma^+_{\boldsymbol{\sigma}'} > 0$ and $\gamma^-_{\boldsymbol{\sigma}'} > 0$ satisfying

$$(3.1)\qquad \begin{array}{l} \sum_{\boldsymbol{\sigma}'} \gamma^+_{\boldsymbol{\sigma}'} = \sum_{\boldsymbol{\sigma}} \gamma^-_{\boldsymbol{\sigma}'} \quad = 1 \\ \left.\begin{array}{ll} \sum_{\boldsymbol{\sigma}'} \lambda_{k,[\boldsymbol{\sigma}',+]}\gamma^+_{\boldsymbol{\sigma}'} & = 0, \\ \sum_{\boldsymbol{\sigma}'} \lambda_{k,[\boldsymbol{\sigma}',-]}\gamma^-_{\boldsymbol{\sigma}'} & = 0, \end{array}\right\} \quad (k = 1, \dots, m). \end{array}$$

Using (2.9) for $k = \hat{m} + 1$, we now have

$$\begin{aligned} \sum_{\boldsymbol{\sigma}'} \lambda_{\hat{m}+1,[\boldsymbol{\sigma}',+]} \gamma^{+}_{\boldsymbol{\sigma}'} &=: u^{+} &< 0, \\ \sum_{\boldsymbol{\sigma}'} \lambda_{\hat{m}+1,[\boldsymbol{\sigma}',-]} \gamma^{-}_{\boldsymbol{\sigma}'} &=: u^{-} &> 0, \end{aligned}$$

and can then define

$$\begin{aligned} &\alpha := \frac{u^{-}}{u^{-} - u^{+}} \in (0,1), && (1-\alpha) = -\frac{u^{+}}{u^{-} - u^{+}} \\ &\gamma_{[\boldsymbol{\sigma}',+]} := \alpha \gamma^{+}_{\boldsymbol{\sigma}'}, && \gamma_{[\boldsymbol{\sigma}',-]} := (1-\alpha)\gamma^{-}_{\boldsymbol{\sigma}'}. \end{aligned} \tag{3.2}$$

It is easy to see that (2.11) is satisfied for $\hat{m} + 1$ by using $\gamma_{[\boldsymbol{\sigma}',\pm]} > 0$ and we have the desired construction for arbitrary m by induction.

We indicate later (cf., Theorem 3) a physical interpretation of the solution constructed in this way, which we denote by $(\hat{\gamma}_{\boldsymbol{\sigma}})$. At this point we do note that, since the $\{f^{\boldsymbol{\sigma}}(\xi)\}$ and $\{\mathbf{n}_k(\xi)\}$ are locally Lipschitzian in their dependence on ξ (we take this as a minimal interpretation of the 'smoothness' originally assumed for the individual vector fields and for the switching surfaces), all the 'λ's appearing in the defining system (2.11) are also locally Lipschitzian. Following the inductive construction above then shows that $\xi \mapsto (\hat{\gamma}_{\boldsymbol{\sigma}})$ is also locally Lipschitzian.

We next note that the homogeneous adjoint system

$$\rho_0 + \sum_{k=1}^{m} \lambda_{k,\boldsymbol{\sigma}} \rho_k = 0 \qquad (\text{all } \boldsymbol{\sigma}) \tag{3.3}$$

has only the trivial solution. [To see this, suppose the contrary and first consider $\boldsymbol{\sigma}$ such that $\sigma_k = \operatorname{sgn} \rho_k$ where $\rho_k \neq 0$, giving $\lambda_{k,\boldsymbol{\sigma}} \rho_k < 0$ for each such term by (2.9) and so $\rho_0 > 0$, and then consider $\boldsymbol{\sigma}' = -\boldsymbol{\sigma}$, now giving $\sum_k \lambda_{k,\boldsymbol{\sigma}'} \rho_k > 0$ and so $\rho_0 < 0$ — a contradiction.] By the Fredholm Alternative, this shows that the $m + 1$ equations in (2.11) are independent — i.e., that $\mathbf{A}$ has full rank $m + 1$. It follows that $dim \mathcal{N}(\mathbf{A}) =: \nu = 2^m - m - 1$, fixing the 'degree of ambiguity' for these coefficients: each $(\gamma_{\boldsymbol{\sigma}}) \in \mathcal{C}$ has a representation

$$(\gamma_{\boldsymbol{\sigma}}) = (\hat{\gamma}_{\boldsymbol{\sigma}}) + \sum_{n=1}^{\nu} c_n \, (\tilde{\gamma}^{n}_{\boldsymbol{\sigma}}) \tag{3.4}$$

where $\{(\tilde{\gamma}^n_{\boldsymbol{\sigma}}) : n = 1, \ldots, \nu\}$ is a basis for $\mathcal{N}(\mathbf{A}(\xi))$. From the (already noted) Lipschitz continuity of the ξ-dependent 'λ's it follows that each $\xi \mapsto (\tilde{\gamma}^n_{\boldsymbol{\sigma}})$ can be taken as locally Lipschitzian. ■

This ambiguity of the coefficients indicates a *potential* ambiguity of the field. Whether such an ambiguity is realized, — i.e., whether $\mathcal{F}(\xi)$ might or might not be a singleton — still depends on the vector fields $\{f^{\boldsymbol{\sigma}}\}$. The determination of $\mathcal{C}_+$ depends only (through the $\{\lambda_{k,\boldsymbol{\sigma}}\}$) on the *transverse* components of these vector fields but, once $(\gamma_{\boldsymbol{\sigma}})$ would be specified (in $\mathcal{C}_+$ — making the transverse component of the resultant field φ of (2.8) vanish — then the sliding mode one obtains depends only on the *parallel* components of $f^{\boldsymbol{\sigma}}$. We turn now to consideration of the set $\mathcal{F}(\xi)$ of Filippov vectors at ξ via (2.14). We see later (Theorem 2) that the ambiguity of the weak formulation "(1.3) with $\mathcal{F}$" is inherent in that any solution is actually *implementable* in a sense to be discussed in the next section. Thus, any further resolution of this ambiguity must rely on the imposition of some additional restriction or on the use of some additional information about the nature of the implementation.

Before proceeding further, we note a continuity result for $\mathcal{F}(\xi)$.

LEMMA 1: *Given $\xi_* \in \mathcal{Z}$, there exists L, ε such that, for all $\xi \in \mathcal{Z}$ with $|\xi - \xi_*| < \varepsilon$ and any $\varphi \in \mathcal{F}(\xi)$, there exists $\varphi_* \in \mathcal{F}(\xi_*)$ with $|\varphi - \varphi_*| \leq L|\xi - \xi_*|$.*

PROOF: For any $\varepsilon > 0$ such that $\mathcal{O}_\varepsilon := \{\xi \in \mathcal{Z} : |\xi - \xi_*| \leq \varepsilon\}$ lies in $\mathcal{O}$, the Lipschitz continuity noted above for (3.4) holds uniformly by compactness and $\xi \mapsto \mathbf{F}(\xi)$ is similarly (uniformly) Lipschitzian here. Using (2.14), (3.4), the given φ has the form:

$$\varphi = \mathbf{F}(\gamma_{\boldsymbol{\sigma}}) \quad \text{with } (\gamma_{\boldsymbol{\sigma}}) = \left[(\hat{\gamma}_{\boldsymbol{\sigma}}) + \sum_n c_n (\tilde{\gamma}^n_{\boldsymbol{\sigma}}) \right].$$

Determine $r \in (0, 1]$ maximal so, when evaluated at ξ_* with the same $\{c_n\}$, the entries of $(\gamma_{\boldsymbol{\sigma}})_* = [(\hat{\gamma}_{\boldsymbol{\sigma}}) + r \sum_n c_n (\tilde{\gamma}^n_{\boldsymbol{\sigma}})]$ are all nonnegative so $(\gamma_{\boldsymbol{\sigma}})_* \in \mathcal{C}_*$; one easily sees that r satisfies a suitable Lipschitz estimate so that, for a constant L depending on the uniform Lipschitz continuity of the various factors, one has the desired estimate for $\varphi_* := \mathbf{F}(\xi_*)(\gamma_{\boldsymbol{\sigma}})_*$. ■

Finally, for consideration of uniqueness we observe that a (trivial) *sufficient* condition would be that the parallel components $(f^{\boldsymbol\sigma})^{\|}$ — i.e., projections of $f^{\boldsymbol\sigma}$ to the tangent space $\mathcal{T}$ — would be independent of $\boldsymbol\sigma$. A rather more interesting sufficient condition is the (local) continuity of the jump across $\mathcal{Z}_k$ for each k (in crossing any other switching surface). It is easily seen that this continuity implies an additive representation with respect to the sensors:

$$f^{\boldsymbol\sigma} = \bar f + \sum_k \sigma_k w_k. \tag{3.5}$$

(so the jump across each $\mathcal{Z}_k$ would be $2w_k$) as a condition on the fields and uniqueness of the sliding mode in this situation[3] was already shown in [2].

4. Sliding modes via 'chattering approximations'

Any model is of course an idealization of some actual physical situation. In the real world, a 'design description' such as (1.1) cannot just happen, but must be *made* to happen — 'realized' using physical sensors to determine the 'truth' of (compound) Boolean propositions "$x(t)$ is in $\mathcal{O}^{\boldsymbol\sigma}$" and actuators

---

[3]For completeness in our present framework, we include here the following:

PROOF: First note that, in the setting of (3.5), the inwardness condition (2.9) implies linear independence of the transverse components $\{w_k^\perp\}$. [To see this, set $\bar\mu_k := \mathbf{n}_k \cdot \bar f$ and $\mu_{j,k} := \mathbf{n}_k \cdot w_j^\perp$ so (2.9) gives, for all $\boldsymbol\sigma$ and each k, that $\operatorname{sgn}\,[\bar\mu_k + \Sigma_j \sigma_j \mu_{j,k}] = -\sigma_k$. For any $\rho \in \mathbb{R}^m$ one can choose $\boldsymbol\sigma$ by $\sigma_k = \operatorname{sgn}\rho_k$ (arbitrary if $\rho_k = 0$) to get

$$\rho_k \bar\mu_k + \sum_j \sigma_j \mu_{j,k}\rho_k < 0 \quad \text{when } \rho_k \neq 0. \tag{3.6}$$

If $\{w_k^\perp\}$ were dependent, then the $m \times m$ matrix $(\mu_{j,k})$ must be singular so we could choose $0 \neq \rho \in \mathbb{R}^m$ as an eigenvector: $\sum_k \mu_{j,k}\rho_k = 0$ so summing (3.6) over k gives $\sum_k \rho_k \bar\mu_k < 0$ (strict inequality). Since $-\rho$ would also be such an eigenvector, this is a contradiction.]

For $(\gamma_{\boldsymbol\sigma}) \in \mathcal{N}(\mathbf{A})$ one would have vanishing of the transverse component $(\sum_{\boldsymbol\sigma} \gamma_{\boldsymbol\sigma} f^{\boldsymbol\sigma})^\perp$ so

$$\begin{aligned} 0 &= \textstyle\sum_{\boldsymbol\sigma} \gamma_{\boldsymbol\sigma} \left(\bar f + \sum_k \sigma_k w_k\right)^\perp \\ &= \left(\textstyle\sum_{\boldsymbol\sigma} \gamma_{\boldsymbol\sigma}\right) \bar f + \sum_{\boldsymbol\sigma} \left(\sum_k \gamma_{\boldsymbol\sigma} \sigma_k w_k^\perp\right) = \sum_k \left(\sum_{\boldsymbol\sigma} \sigma_k \gamma_{\boldsymbol\sigma}\right) w_k^\perp. \end{aligned}$$

By the independence, each of these coefficients $\sum_{\boldsymbol\sigma} \sigma_k \gamma_{\boldsymbol\sigma}$ must vanish so the tangential component $(\sum_{\boldsymbol\sigma} \gamma_{\boldsymbol\sigma} f^{\boldsymbol\sigma})^{\|} = \sum_k (\sum_{\boldsymbol\sigma} \sigma_k \gamma_{\boldsymbol\sigma})\, w_k^{\|}$ also vanishes whence $(\gamma_{\boldsymbol\sigma}) \in \mathcal{N}(\mathbf{F})$. Thus, $\mathcal{N}(\mathbf{A}) \subset \mathcal{N}(\mathbf{F})$ and, by (2.16), one has the desired uniqueness. ∎

to provide the relevant mode $[\dot{x} = f^{\boldsymbol{\sigma}}]$. These sensors and actuators cannot actually operate instantaneously, but can only be assumed to act on a faster time scale than the 'natural' time scale for consideration of (1.1) so as to be apparently discontinuous despite the principle *"Natura in operationibus non facit saltus."* In any actual situation, the system (1.1) does not respond instantaneously to the sensor functionals but rather the response takes place in some neighborhood of the switching surfaces $\mathcal{Z}_k$.

The conceptually simplest notion of 'implementation' for (1.1) is by 'open loop' approximation: Suppose for (small) $\varepsilon > 0$ there is a neighborhood $\mathcal{N} = \mathcal{N}_\varepsilon$ of $\mathcal{Z}$ within which all the relevant modes $\{f^{\boldsymbol{\sigma}}\}$ are 'available'. It would then be possible to select modes $f^{\boldsymbol{\sigma}}$ arbitrarily (on very short time intervals, corresponding to a fast time scale) provided we remain within $\mathcal{N}_\varepsilon$. Any such sequential selection produces a zigzag trajectory x_ε which we call a 'chattering approximation' subject to the constraint that we must remain within $\mathcal{N}_\varepsilon$. This trajectory x_ε behaves, in some sense, like a sliding mode and, indeed, we may think of a limiting process with the constraining neighborhood $\mathcal{N}_\varepsilon$ shrinking down to $\mathcal{Z}$ as $\varepsilon \to 0$ — with a necessarily increasingly rapid time scale for the selections. It may then happen that $x_\varepsilon \to \bar{x}$ (uniformly on some time interval) and it is then easy to see that this limit trajectory $\bar{x}$ is a solution of (1.3) — i.e., a sliding mode. In this case we view the ε-solutions x_ε as approximate implementations of $\bar{x}$ and, *in this sense,* say that the sliding mode $\bar{x}$ is *realizable.*

THEOREM 2: *Every Filippov solution is realizable in the sense above.*

PROOF: We wish to show that an arbitrary solution $\bar{x}(\cdot)$ of (1.3) can be uniformly approximated on an interval — which, without loss of generality, we may take to be $[0, 1]$ for expository convenience — by an ε-solution $\tilde{x}(\cdot)$ of the special form: $\tilde{x}^{\,\cdot} = \tilde{\varphi}$ where $\tilde{\varphi}$ is to be piecewise constant with values $\tilde{\varphi}(t) = f_j(\xi)$ at each t, such that ξ and $\tilde{x}(t)$ are uniformly close.

Let $\mathcal{N} \subset \mathcal{O}$ be a compact (tubular) neighborhood of the trajectory $\{\bar{x}(t) : t \in [0, 1]\}$; by compactness we have on $\mathcal{N}$ a uniform bound M and Lipschitz constant L for the 2^m vector fields $\{f_j\}$. Note that when $|\xi - \xi_*| < \delta$ one has

$$\mathcal{H}_\varepsilon(\xi) = \overline{hull}\left\{f_j(\tilde{\xi}) : |\tilde{\xi} - \xi| < \varepsilon\right\} + \mathcal{B}_\varepsilon \subset \overline{hull}\ \{f_j(\xi_*)\} + \mathcal{B}_{\varepsilon+(\varepsilon+\delta)L}.$$

Setting $t_n := n\delta$ with $\delta = 1/N$ $(n = 1, \ldots, N)$, we then have, from (1.3),

$$(4.1) \qquad \begin{aligned} &[\bar{x}(t_{n+1}) - \bar{x}(t_n)] = \int_{t_n}^{t_n+\delta} \bar{\varphi}(s)\, ds = \hat{\varphi}_n + r_n \\ &\hat{\varphi}_n := \sum_j c_j^n f_j(\bar{x}(t_n)), \quad |r_n| \leq \rho \ \ (\text{ any } \rho > L\delta \ ) \end{aligned}$$

since $\bar{\varphi}(s) \in \mathcal{H}_\varepsilon(\bar{x}(s)) \subset \overline{hull}\,\{f_j(\bar{x}(t_n))\} + \mathcal{B}_\rho$ on $[t_n, t_{n+1}]$ with $\rho = \varepsilon + (\varepsilon + \delta)L$ for some $\varepsilon > 0$. The Filippov ε-solution $\hat{x}(\cdot)$ defined by $\hat{x}^{\,\cdot} = \hat{\varphi}_n$ on $[t_n, t_{n+1}]$ satisfies $|\hat{x} - \bar{x}| \leq 2\rho \approx 2L/N$ on $[0, 1]$ and we may make this arbitrarily small (so, in particular, $\hat{x}$ remains in the interior of $\mathcal{N}$) by taking N large enough.

Further subdividing the intervals $[t_n, t_{n+1}]$ into N' equal parts and then each of those into 2^m parts proportional to the coefficients c_j^n, we may replace the convex combination defining $\hat{\varphi}_n$ by the successive use of the individual vectors[4] $f_j^n = f_j(\bar{x}(t_n))$ with the resulting $\tilde{x}$ arbitrarily close to $\hat{x}$ (and so in $\mathcal{N}$ and close to $\bar{x}$) if N' is taken large enough. ■

While we can produce (as an approximation) any solution of (1.3) by a control strategy teleologically designed to do precisely this, we are more interested in the operation of some more natural implementations of (1.1) and the associated sliding modes. We think of a 'natural implementation' or mechanism as a scheme, parametrized by $\eta \to 0$, which suitably produces $\dot{x}_\varepsilon =: \varphi$ as a control law in a way (meeting the constraint $x_\varepsilon(t) \in \mathcal{N}_\varepsilon$) which is autonomous and causal — indeed, which depends only on (the state and) the scaled sensor values $\hat{y}_k = Y_k(x)/\eta_k$. It is difficult to describe formally what we mean here by 'natural', but some examples of such implementations are:

forward Euler retaining a constant value $f^{\boldsymbol{\sigma}(x)}$ for a time step $\eta_k\delta$ (e.g., without switching if one does land exactly on some $\mathcal{Z}_k$).

hysteresis retaining a constant mode $f^{\boldsymbol{\sigma}}$ until reaching the boundary of $\mathcal{N}_\varepsilon := \{x : |\hat{y}_k| < 1\}$ and then switching the component σ_k associated with the active boundary face; for some partial results, see [6].

[4]It is easy to see that the alternate interpretation of simple modal selection in (1.1), using $f_j(\tilde{x}(\cdot))$ rather than $f_j(\bar{x}(t_n))$ on each of these finest intervals, would induce a modification of the trajectory but this further perturbation would also vanish in the limit.

sigmoid blending blending (interpolating) the fields in a way corresponding to the values $\hat{y}_k$; see the next section.

stochastic switching using $f^{\boldsymbol{\sigma}}$ with independent switching of each σ_k taken as a Poisson process with parametrization $\lambda(|\hat{y}_k|)$, adjusted to switch with certainty by $|\hat{y}_k| = 1$.

stochastic equation adding (colored[5]) Brownian noise, replacing the deterministic ordinary differential equation (1.1) by a stochastic equation (essentially of Ito type)

$$dx = f^{\boldsymbol{\sigma}}(x)\,dt + \eta A\,dw \tag{4.2}$$

with $\boldsymbol{\sigma} = \boldsymbol{\sigma}(x)$, noting that one almost never has $x(t) \in \mathcal{Z}_k$ except on a nullset; see [1].

Observe that each of these mechanisms is scaled essentially by the dependence on $\hat{y} := Y_k(x)/\eta_k$. Effectively, each factor η_k may be interpreted as a 'sensitivity' or 'reaction speed' of the k^{th} sensor functional — keeping the nominal functional $Y_k(\cdot)$ fixed for convenience. In any of these implementations the switching acts on a time scale — determined by η_k — which is much faster than our consideration of (1.1) so we may analyze the relation of the corresponding chattering trajectory to a putative sliding mode by a limiting procedure as for Theorem 2 with each $\eta_k \to 0$. Such an analysis then assumes introduction of an intermediate time scale (corresponding roughly to the subdivision by $\{t_n\}$ in the proof of Theorem 2, above) fast enough that $x(\cdot)$ changes little on this scale so the analysis can remain local yet is slow enough in comparison to the fast scale of mode switching that it corresponds there to asymptotically long times. Note that the implication of the localization given by the intermediate time scale is that there is no loss of generality in using the coordinatization given by (2.4) and (2.5) with an assumption of constant vector fields $f^{\boldsymbol{\sigma}}$ when looking at fast time scales; henceforth we restrict our attention to that setting.

There are now two[6] interesting possibilities: either one has a hierarchy

[5] The matrix $A = A(x)$ in (4.2) is necessarily included to allow for a transformation such as (2.5); it is also possible to consider $A = A^{\boldsymbol{\sigma}(x)}$.

[6] The mixed case could easily be analyzed by combining the analyses of these two possibilities. There would also be no difficulty in having r_k defined as a limit in the second case.

of distinct time scales (so ratios $\eta_k/\eta_{k'} \to 0, \infty$) or there is a single fast time scale (i.e., $\eta_k = r_k \eta$ with r_k fixed as $\eta \to 0$).

The first of these possibilities was considered in [5] in the context of hysteresis: the ambiguity of Theorem 1 is now removed, and a unique sliding mode results, corresponding to the specific coefficient construction for $(\hat{\gamma}_{\boldsymbol{\sigma}})$. We note here that the argument of [5] generalizes to all our 'natural implementations'.

THEOREM 3: *In the case of hierarchical 'time scale separation' there is a uniquely determined sliding mode $\dot{x} = \sum_{\boldsymbol{\sigma}} \hat{\gamma}_{\boldsymbol{\sigma}} f^{\boldsymbol{\sigma}}$, given by the recursive construction of $(\hat{\gamma}_{\boldsymbol{\sigma}})$ via (3.2), etc., in the proof of Theorem 1.*

PROOF: With no loss of generality, we assume the sensors are indexed so the scale separation has $\eta_1 \to 0$ fastest, etc.

Since the neighborhood $\mathcal{N}$ is extremely narrow — i.e., $\mathcal{O}(\eta_1)$ — in the y_1-direction, the time scale for consideration of interaction with the switching surface $\mathcal{Z}_1$ is $\tau := t/\eta_1$ and we observe that if we were also to scale spatially by η_1 (so the vector fields remain unchanged), then $|\hat{y}_2|, \ldots, |\hat{y}_m|$ each remains 'large'. We see that the determination, by our implementation, of the field to be used becomes independent of $\{\hat{y}_k : k > 1\}$ for long times on this fastest time scale and so effectively depends only on $\hat{y}_1$ — i.e., effectively we have $m = 1$ for analysis on this time scale with zigzagging across $\mathcal{Z}_1$ (and asymptotically slow variation of $\hat{y}_2, \ldots, \hat{y}_m$ as well as of z) using only the two relevant fields — here temporarily denoted by $F^{\pm}$ — corresponding to $\sigma_1 = \operatorname{sgn} \hat{y}_1 = \pm$ with $\hat{y}_k \to \pm\infty$ for $k > 1$.

We already know the correct analysis for this situation: for $m = 1$ there is a unique sliding mode (here lying in $\mathcal{Z}_1$) and the chattering approximation on this fastest time scale necessarily approximates that with long term averages (on this scale) using $F^{\pm}$ the appropriate fractions of the time to keep $\hat{y}_1$ bounded — i.e., the averaged effect is

$$
\begin{aligned}
&F_1 := [\alpha F^+ + (1-\alpha) F^-] \quad \text{with} \\
&\alpha := \frac{u_1^-}{u_1^- - u_1^+}, \quad (1-\alpha) := \frac{-u_1^+}{u_1^- - u_1^+}
\end{aligned}
\tag{4.3}
$$

where the scalars $u_1^{\pm}$ are the y_1-components of $F^{\pm}$ and we note that $u_1^+ < 0 < u_1^-$ by the inwardness condition. Clearly this F_1 still depends on the particular $\boldsymbol{\sigma}' = [\sigma_2, \ldots, \sigma_m]$.

Now, looking at the next fastest time scale (corresponding to η_2), we proceed similarly — using the sliding modes $F_1^{\boldsymbol{\sigma}'}$ we have just found in $\mathcal{Z}_1$ since, by our prior analysis, we know that this defines the appropriate dynamics on time scales slower than that of η_1. As before, the scaled sensor values $\hat{y}_k$ ($k > 2$) retain their signs while $\sigma_2 := \operatorname{sgn} \hat{y}_2$ switches. On this η_2 time scale we again have a problem (within $\mathcal{Z}_1$) with zigzagging across the single switching surface $\mathcal{Z}_2$ and a similar analysis applies to get once more a unique sliding mode in $\mathcal{Z}_1 \cap \mathcal{Z}_2$.

Proceeding recursively in this fashion — effectively having $m = 1$ at each of the distinct time scales — we construct a unique sliding mode in $\mathcal{Z}$. Comparing this construction through (4.3) to that through (3.2) in the proof of Theorem 1, we see that we have obtained precisely the same sliding mode $(\hat{\gamma}_{\boldsymbol{\sigma}})$ as there. Thus, one physical interpretation of that construction is as here: the use of (any) natural implementation with widely separated sensor sensitivities. ■

For the remainder of the present paper, we restrict our attention to the second of the possibilities mentioned and assume the scalings are all commensurate with each other. Noting that the ratios r_k can be absorbed by an initial normalizing replacement $Y_k \leftarrow Y_k/r_k$ so one has only a single $\eta \to 0$, we henceforth use the coordinatization (2.4) with this normalization so $\eta_k = \eta$ for each k. As noted earlier, the existence of an intermediate time scale means that we may take each $f^{\boldsymbol{\sigma}}$ as constant: its value at ξ_*, which in the new coordinatization is taken to be the origin. Suppose this intermediate time scale is given by taking $\tau := t/\varepsilon$ (so the fast time variable is $s := t/\eta = \tau/\delta$ where both $\varepsilon \to 0$ and $\delta := \eta/\varepsilon \to 0$) and we also rescale space by setting $\omega := w/\varepsilon$ so fields are unchanged. Using the representation $\varphi(s) = \sum_{\boldsymbol{\sigma}} c_{\boldsymbol{\sigma}}(s) f^{\boldsymbol{\sigma}}$ for the fast time scale, we then have

$$\begin{aligned}
\frac{\omega(h) - \omega(0)}{h} &= \frac{1}{h}\int_0^h \hat{\varphi}(\tau)\,d\tau = \frac{1}{S}\int_0^S \varphi(s)\,ds \\
&= \frac{1}{S}\int_0^S \sum_{\boldsymbol{\sigma}} c_{\boldsymbol{\sigma}}(s) f^{\boldsymbol{\sigma}}\,ds \\
&= \sum_{\boldsymbol{\sigma}} \left[\frac{1}{S}\int_0^S c_{\boldsymbol{\sigma}}(s)\,ds\right] f^{\boldsymbol{\sigma}}
\end{aligned}$$

with $S := h/\delta \to \infty$ — i.e., the coefficient $\gamma_{\boldsymbol{\sigma}}$ of each field $f^{\boldsymbol{\sigma}}$ in the representation (2.8) for the resulting sliding mode is just the asymptotic time average (for the fast time scale) of its use: $\gamma_{\boldsymbol{\sigma}} = \gamma^*_{\boldsymbol{\sigma}}$ given by

$$\gamma^*_{\boldsymbol{\sigma}} := \lim_{S\to\infty} \frac{1}{S} \int_0^S c_{\boldsymbol{\sigma}}(s)\, ds \tag{4.4}$$

— *provided, of course,that this limit exists, preferably independent of the precise initial data.* From our previous analysis we know that this must give $(\gamma^*_{\boldsymbol{\sigma}}) \in \mathcal{C}_+(\xi_*)$. At this point we also emphasize that the nature and scaling of our 'natural mechanisms' are such that:

(1) the dynamics at the fast time scale are entirely independent of the scale parameter η;

(2) the dynamics depend only on the $\{\hat{y}_k\}$ so one may project to the transverse component and consider the projected dynamics in $\mathbb{R}^m$.

(3) we could 'freeze' the vector fields $f^{\boldsymbol{\sigma}}$ as constants (evaluating at $(0,z)$) for the analysis leading to (4.4) on the fast time scale but for use of this when we return to the original (natural) time scale, we must restore the dependence on $z \in \mathcal{Z}$ in determining $(\gamma^*_{\boldsymbol{\sigma}}(z))$ and so the sliding mode

$$f^* = f^*(z) = \sum_{\boldsymbol{\sigma}} \gamma^*_{\boldsymbol{\sigma}}(z)\, [f^{\boldsymbol{\sigma}}(z)]^{\|}\,. \tag{4.5}$$

As is consistent with Theorem 3, we observe that any change in the relative sensitivites of the sensors necessarily affects the resulting sliding mode by affecting the transformed fields $f^{\boldsymbol{\sigma}}$ determined by the normalization we have imposed.

A sufficient condition for existence of the limits in (4.4) is that a stable stationary point be a global attractor for the fast dynamics so, asymptotically, the coefficients $c_{\boldsymbol{\sigma}}$ become constant. This is possible — indeed, our main result is that it always holds for the blending mechanism when $m = 1, 2$ — but, more generally, we note that we need only find a global attractor with a suitable invariant measure. Looking again at our scaling, what we have done is to consider dynamics $[y, z]^{\cdot} = \hat{f}(y/\eta, z)$ — where, symbolically, $\hat{f}$ is associated with the particular 'natural implementation' we might be using — and, noting that (setting $\hat{y} := y/\eta$) this is equivalent to the singularly perturbed system

$$\begin{aligned} \eta\hat{y}^{\,\cdot} &= \left[\hat{f}(\hat{y}, z)\right]^{\perp} \\ \dot{z} &= \left[\hat{f}(\hat{y}, z)\right]^{\|} \end{aligned} \tag{4.6}$$

for which, under the appropriate conditions, singular perturbation theory gives a sliding mode — $\dot{z} = f^*(z)$, using (4.5) — as the reduced dynamics.

5. Blending

We turn to consideration of a specific mechanism, *sigmoid blending*, for which we show (for the case $m = 2$) that the fast dynamics described in the preceding section has a stationary point as global attractor, leading to a well-defined and easily computable sliding mode, independent of the details of the blending. The known case $m = 1$ is treated as illustratory of the basic ideas.

The point of blending is the following. As remarked earlier, in any actual mechanism the system cannot switch instantaneously and discontinuously between the different fields on either side of a switching surface $\mathcal{Z}_k$. We now assume there is some small region near $\mathcal{Z}_k$ in which one interpolates between these vector fields in some continuous convex manner to produce a 'blended vector field' which is what is then actually used[7] for the dynamics within the region.

More precisely, we begin with some box $\mathcal{B}$ in $\mathbb{R}^m$ (say, e.g., $\mathcal{B} = [-K, K]^m$) with vertices $y^{\boldsymbol{\sigma}}$ in the orthants $\mathcal{O}^{\boldsymbol{\sigma}}$ and assume we have (smooth) coefficient maps: $y \mapsto \gamma_{\boldsymbol{\sigma}}(y) : \mathcal{B} \to [0,1]$ satisfying $\sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}}(y) \equiv 1$; assume further that these $\{\gamma_{\boldsymbol{\sigma}}\}$ are such that on each face of $\mathcal{B}$ all the $\gamma_{\boldsymbol{\sigma}}$ vanish for which $\boldsymbol{\sigma}$ is associated with the opposite face. Now we define a smooth vector field $\hat{f}$ on $\mathcal{B}$ by interpolation:

$$\hat{f}(y) := \sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}}(y) f^{\boldsymbol{\sigma}} \qquad \text{for } y \in \mathcal{B} \subset \mathbb{R}^m \tag{5.1}$$

and approximate (1.1) there by

$$\dot{x} = \hat{f}(y/rsize) \qquad \text{for } x = (y, z). \tag{5.2}$$

Since $\dot{x}$ only depends on the transverse y-component of x, we may project and scale to consider separately the dynamics for $y \hookleftarrow y/\eta$ on the the fast

[7] An alternate interpretation would be to use the vector fields $f^{\boldsymbol{\sigma}}$ individually in sequential rotation on a still more rapid time scale (like the fastest scale in the proof of Theorem 2) so that each occurs in isolation on that scale but altogether uses a fraction of the time, locally on what we *had* been considering the 'fast' time scale, corresponding to its coefficient in the convex combination.

time scale $t \leftarrow t/\eta$, given by

$$\dot{y} = \hat{f}^*(y) := \left[\hat{f}(y)\right]^\perp = \sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}}(y)\,[f^{\boldsymbol{\sigma}}]^\perp\,. \tag{5.3}$$

We easily see that the condition on $\{\gamma_{\boldsymbol{\sigma}}\}$ at the faces together with the inwardness condition (2.9) imply that the vector field $\hat{f}^*$ points strictly inward at each point of $\partial\mathcal{B}$ so $\mathcal{B}$ is necessarily an invariant set for the flow (5.3) — whence the x-flow remains in an arbitrarily small neighborhood of $\mathcal{Z}$ provided η is taken small enough.

For future reference, we also note that a simple degree argument (noting that inwardness at $\partial\mathcal{B}$ implies that $\hat{f}^*$ has Hopf index 1) shows that there must be some point y^* in the interior of $\mathcal{B}$ at which $\hat{f}^*(y^*) = 0$, i.e., $\sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}}(y^*) f^{\boldsymbol{\sigma}} = 0$ so $(\gamma^*_{\boldsymbol{\sigma}}) := (\gamma_{\boldsymbol{\sigma}}(y^*)) \in \mathcal{C}_+$. This y^* is, of course, a stationary point of the flow (5.3). If it would happen that $\{y^*\}$ were a global attractor for this flow, then (as was noted from (4.4) at the end of the previous section)

$$f^* := f(y^*) = \sum_{\boldsymbol{\sigma}} \gamma^*_{\boldsymbol{\sigma}}\,[f^{\boldsymbol{\sigma}}]^{\|} \tag{5.4}$$

is the uniquely determined sliding mode.

Sigmoid blending is a particular case of this general 'blending mechanism', constructing the interpolation in a special product form, related to 'fuzzy logic'. A sigmoid function is a nondecreasing function $\alpha : \mathbb{R} \to [0,1]$ with $\alpha(r) \to 0, 1$ as $r \to -\infty, +\infty$; for convenience, we assume here that any sigmoid function $\alpha(\cdot)$ we consider is smooth with $\alpha'(r) > 0$ when $\alpha(r) \in (0,1)$ and that there is some K such that $\alpha(-K) = 0$, $\alpha(K) = 1$. One interpretation of this $\alpha(r)$ is as the truth value, in fuzzy logic, of the elementary proposition "$[r > 0]$".

The 'fuzzified' truth value of a logical proposition such as "$\hat{y}$ is positive" is given by a function $\alpha(\cdot) : \mathbb{R} \to [0,1]$ for which $\alpha(y) = 0$ when $\hat{y}$ is 'very negative' — say, when $\hat{y} \le -K$ — and $\alpha(\hat{y}) = 1$ when $\hat{y}$ is 'very positive' — say, when $\hat{y} \ge +K$ — with intermediate values, 'partial truth', for smaller $|\hat{y}|$, noting that our scaling is such that this corresponds to having x close to the (single) switching surface $\mathcal{Z}$. Note that $(1-\alpha)$ is then the 'fuzzified' truth value of the negation of this proposition. We use this same α to interpolate between the adjacent vector fields $f^\pm$ to get $\varphi = \alpha(y)f^+ + (1-\alpha(y))f^-$ — in some sense a 'fuzzification' of the assertion that

"$\varphi := \dot{x}$ is f^{+} for $y := Y(x)$ positive and f^{-} for y negative."

For $m > 1$, our treatment is precisely the standard 'fuzzification' of a conjunctive logical form

$$[\xi \in \mathcal{O}^{+-\cdots}] \equiv [Y_1(\xi) > 0] \wedge \neg[Y_2(\xi) > 0] \wedge \cdots \tag{5.5}$$

from its elements — i.e., the 'fuzzified' truth value of "ξ is in $\mathcal{O}^{+-\cdots}$" would be the product of the 'fuzzified' truth values of the components "y_1 is positive", $\neg$"y_2 is positive", etc., and we would use this product as the coefficient of $f^{+-\cdots}$ in our interpolation.

Given $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_m) \in [0,1]^m$ (e.g., as a column vector) and $\boldsymbol{\sigma} := (\sigma_1, \ldots, \sigma_m) \in \{\pm\}^m$, we set

$$\gamma_{\boldsymbol{\sigma}}(\boldsymbol{\alpha}) := \prod_{k=1}^{m} \alpha_k^{\sigma_k} \text{ with } \alpha_k^{\sigma_k} := \begin{cases} \alpha_k(y), & \text{if } \sigma_k = `-\text{'}, \\ 1 - \alpha_k(y), & \text{if } \sigma_k = `+\text{'}. \end{cases} \tag{5.6}$$

This construction gives a smooth m-dimensional manifold

$$\Gamma := \{(\gamma_{\boldsymbol{\sigma}}(\boldsymbol{\alpha})) : \boldsymbol{\alpha} \in [0,1]^m\} \subset [0,1]^{2^m}. \tag{5.7}$$

Noting that (5.6) gives $\sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}} = 1$, we can use these coefficients for a convex combination, as before

$$\varphi_* = \varphi_*(\boldsymbol{\alpha}) := \sum_{\boldsymbol{\sigma}} \gamma_{\boldsymbol{\sigma}}(\boldsymbol{\alpha})\, f^{\boldsymbol{\sigma}}. \tag{5.8}$$

Of course, we want the *blended field* $\hat{f}$ to interpolate in terms of the transverse state component y. If we have specified m blending functions $\alpha_k(\cdot)$ for $k = 1, \ldots, m$ (i.e., independently for each coordinate variable y_k), then for $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ we can set

$$\begin{gathered} \hat{f} = \hat{f}(y) := \textstyle\sum_{\boldsymbol{\sigma}} c_{\boldsymbol{\sigma}}(y)\, f^{\boldsymbol{\sigma}} \text{ with} \\ c_{\boldsymbol{\sigma}}(y) := \gamma_{\boldsymbol{\sigma}}(\boldsymbol{\alpha}(y)), \quad \boldsymbol{\alpha}(y) := (\alpha_1(y_1), \ldots, \alpha_m(y_1)). \end{gathered} \tag{5.9}$$

Note that in (5.8) and (5.9) we have taken each $f^{\boldsymbol{\sigma}}$ as a constant vector satisfying (2.9), in keeping with our previous general observation about the implication of an intermediate time scale.

Associated with this blended vector field we then have dynamics: $\dot{x} = \varphi$ and, noting that φ depends (through $\boldsymbol{\alpha}$) only on the transverse y-component of x, we may consider separately the y-dynamics:

$$\begin{aligned} &\dot{y} = f^*(y) \quad \text{with} \\ (5.10) \qquad &f^*(y) := [\varphi(y)]^{\perp} = \textstyle\sum_{\boldsymbol{\sigma}} c_{\boldsymbol{\sigma}}(y) u^{\boldsymbol{\sigma}} \\ &\text{where} \qquad f^{\boldsymbol{\sigma}} =: [u^{\boldsymbol{\sigma}}, w^{\boldsymbol{\sigma}}] \end{aligned}$$

while temporarily ignoring the z-component parallel to $\mathcal{Z}$.

For illustration, consider first the case of one switching surface. This case is well-known, but in this context, we explain the notation and general construction, and also we show the use of blending functions leads to the usual result of Filippov.

We suppose we have already coordinatized as in (2.4) so that the switching surface is the coordinate hyperplane $\{y = 0\}$ and y is the *control variable* associated to the switching mechanism. Let the vector fields $f^{(\pm)}$ be denoted f_1, f_2, respectively. As noted in the previous section, we may suppose these fields are constant. We suppose they are decomposed as

$$(5.11) \qquad f_i = \big(u_i, w_i\big), \qquad i = 1, 2$$

on the sets $\{y \geq 0\}$, $\{y \leq 0\}$, respectively. That is, u_i is the component of f_i in the y-direction and w_i is the orthogonal component, parallel to $\mathcal{Z}$. Letting $\alpha(\cdot)$ be a sigmoid blending function, the blended field is then

$$(5.12) \qquad f = \alpha(y) f_1 + \big(1 - \alpha(y)\big) f_2$$

and we clearly obtain the simple y-dynamics:

$$(5.13) \qquad \dot{y} = u(y) = \alpha(y) u_1 + \big(1 - \alpha(y)\big) u_2.$$

Since (2.9) gives $u_1 < 0 < u_2$, we see that $u(y) > 0$ when y is sufficiently negative and $u(y) < 0$ for large enough y with some interval $[a, b]$ on which $u \equiv 0$ ($a = b$ if $u(\cdot)$ is strictly decreasing there). ¿From (5.13) it is clear that $u(y) = 0$ if and only if $\alpha(y) = \alpha^*$, given by

$$(5.14) \qquad \alpha^* := \frac{u_2}{u_2 - u_1} \in (0, 1).$$

For initial data below a we have $y < a$ and $u > 0$ for all t so $y(t) \to a$ as $t \to \infty$ and $\alpha(y(t)) \to \alpha(a) = \alpha^*$; similarly, for initial data above b we have $y(t) \to b$ and $\alpha(y(t)) \to \alpha(b) = \alpha^*$ while data in $[a, b]$ gives a stationary solution so $\alpha(y(t)) \equiv \alpha^*$. Thus, in any case, we have from (5.12) that $f \to f^*$ as $t \to \infty$ with

$$(5.15) \qquad f^* := \alpha^* f_1 + (1 - \alpha^*) f_2 = \left(0, \frac{u_2 w_1 - u_1 w_2}{u_2 - u_1}\right),$$

where we have used (5.14) and have noted that it makes the u-component of f^* vanish.

As with the argument leading to (4.5), this ensures that there is a sliding mode, given by (5.15); we must 'unfreeze' the dependence of the fields on $z \in \mathcal{Z}$ to get the sliding mode dynamics $\dot{z} = f^*(z)$ on the natural time scale. For this case it would also be fairly easy to see the equivalence of this to the analysis of (4.6) by singular perturbation theory — now obtaining $\hat{f}$ for (4.6) as in (5.12), but retaining the dependence on (y,z) of the vector fields f_1, f_2.

We observe that the sliding mode given by (5.15) is explicitly expressed in terms of the fields themselves, independent of any choice of the blending function to be used for the implementation; below we see this again for the case $m = 2$ and note that, when we would have a globally attracting stationary point, simply a consequence of our construction of the blended field through $\boldsymbol{\alpha}$. One also easily sees that (5.15) agrees here — as it must — with the standard Filippov result for one switching surface.

We now proceed to our 'main result', considering the case of two intersecting switching surfaces when using the sigmoid blending mechanism.

It is convenient to assume a coordinatization — say, following (2.4) — (x, y, z) for the state and (u, v, w) for vector fields so the scalars x, y are 'transverse' coordinates, governing the switching, and the scalars u, v are corresponding field components, while the vector z is the 'parallel' coordinate with corresponding vector field component w. Rather than using the previous general notation '$f^{\boldsymbol{\sigma}}$', we now denote the relevant fields by $f_j = (u_j, v_j, w_j$ with indexing corresponding to that of the quadrants of $\mathbb{R}^2$ (i.e., $j = 1$ for $x, y > 0$; $j = 2$ for $x < 0 < y$; $j = 3$ for $x, y < 0$; $j = 4$ for $y < 0 < x$); in this notation, (2.9) becomes

$$u_1,\ u_4 < 0 < u_2,\ u_3 \quad \text{and} \quad v_1,\ v_2 < 0 < v_3,\ v_4\,. \tag{5.16}$$

The blending functions used for x, y are now denoted by $\alpha(\cdot)$, $\beta(\cdot)$, respectively, so — as in (5.8) — the blended field $f = (u, v, w)$ (without subscripts) is now given by

$$\begin{aligned} f = f(x,y) = \ & \alpha(x)\beta(y)f_1 + (1-\alpha(x))\beta(y)f_2 \\ & + (1-\alpha(x))(1-\beta(y))f_3 + \alpha(x)(1-\beta(y))f_4 \end{aligned} \tag{5.17}$$

and, in this notation, (5.3) becomes

$$\dot{x} = u, \quad \dot{y} = v \tag{5.18}$$

where it is convenient to write

$$u(\alpha,\beta) = \alpha\beta u_1 + (1-\alpha)\beta u_2 + (1-\alpha)(1-\beta)u_3 + \alpha(1-\beta)u_4,$$
$$v(\alpha,\beta) = \alpha\beta v_1 + (1-\alpha)\beta v_2 + (1-\alpha)(1-\beta)v_3 + \alpha(1-\beta)v_4, \tag{5.19}$$

— i.e., to express u, v (the x- and y-components, respectively, of f) in terms of α, β, remembering that $\alpha = \alpha(x)$, $\beta = \beta(y)$ for use in (5.18). If we are following this flow, then $\varphi(t) := f(x(t), y(t))$ is given in our present notation by

$$\begin{aligned} \varphi(t) &= \textstyle\sum_{k=1}^{4} c_k(t)\, f_k \quad \text{with} \\ c_k(t) &:= \begin{cases} \alpha(x(t))\beta(y(t)) & \text{for } k=1, \\ (1-\alpha(x(t)))\beta(y(t)) & \text{for } k=2, \\ (1-\alpha(x(t)))(1-\beta(y(t))) & \text{for } k=3, \\ \alpha(x(t))(1-\beta(y(t))) & \text{for } k=4. \end{cases} \end{aligned} \tag{5.20}$$

THEOREM 4: *For sigmoid blending in the case of the intersection of two switching surfaces ($m = 2$), the stationary point for the projected transverse dynamics (5.18) is unique and is a global attractor. Correspondingly, $\Gamma \bigcap \mathcal{C}_+$ is a singleton $\{(\gamma^*_{\boldsymbol{\sigma}})\}$, necessarily independent of the particular choices of the blending functions involved, and gives a well-defined and easily computable sliding mode $\mathbf{F}\ (\gamma^*_{\boldsymbol{\sigma}})$, depending smoothly on the vector fields and on the point $z \in \mathcal{Z}$.*

PROOF: A stationary point (x^*, y^*) for (5.18) is a point at which $u = v = 0$ and it is convenient to examine this, using (5.19), as a quadratic system in the α,β-plane. Note first that the set $\mathcal{C}_v := \{v = 0\}$ is a curve given explicitly by solving the equation $v = 0$ for β in terms of α to get

$$\beta = \hat{\beta}(\alpha) := \frac{\alpha v_4 + (1-\alpha)v_3}{[\alpha v_4 + (1-\alpha)v_3] - [\alpha v_1 + (1-\alpha)v_2]} \tag{5.21}$$

and that $\mathcal{C}_u := \{u = 0\}$ is a curve given similarly by $\alpha = \hat{\alpha}(\beta)$. A stationary point for (5.18) corresponds to a point (α^*, β^*) at which $\mathcal{C}_u$ and $\mathcal{C}_v$ intersect within $[0,1] \times [0,1]$. Observe that, from (5.16), we have $\hat{\alpha}, \hat{\beta} : [0,1] \to (0,1)$ smoothly and the loci $\mathcal{C}_u$, $\mathcal{C}_v$ are hyperbolae having both horizontal and vertical asymptotes (possibly degenerating to the asymptotes if $u_1 - u_2 + u_3 - u_4 = 0$ and/or $v_1 - v_2 + v_3 - v_4 = 0$).

Let us substitute (5.21) into (5.19) to define $\alpha \mapsto \omega(\alpha) := u(\alpha, \hat{\beta}(\alpha)$ for $\alpha \in [0,1]$ so each zero of $\omega(\cdot)$ corresponds to an intersection of $\mathcal{C}_u$, $\mathcal{C}_v$

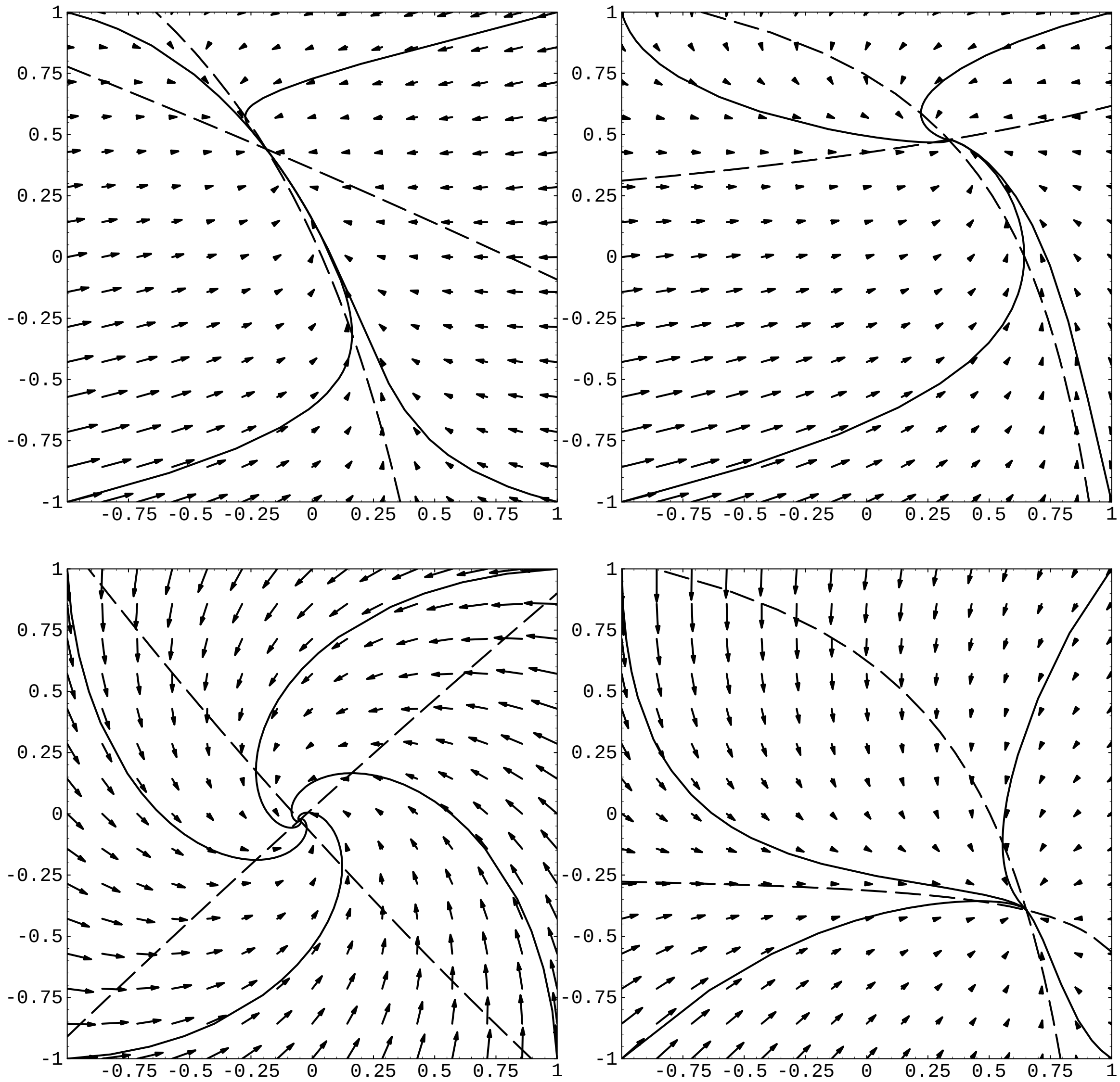

Figure 1. Four examples of flow of (5.19) for differing vectors (u_1, v_1), (u_2, v_2), (u_3, v_3), (u_4, v_4) In this case, we let $\alpha(u) = u$ and $\beta(v) = v$. The general case is qualitatively similar. Shown over the square $-1 \leq u \leq 1$, $1 \leq v \leq 1$ are the vector field, the u and v nullclines (dashed lines) intersecting at the (unique) stationary point, and four orbits with initial conditions at the corners. Note that the nullclines are monotonic and connect opposite sides of square. All orbits converge to the stationary point. The stationary point can be either a node or a spiral, as in the lower left. In the latter case, the system spirals into the stationary point.

within $[0,1]\times[0,1]$ — a finite number, counting multiplicities, since ω is a rational function. We note that $\omega(0) = u(0,\beta_0) > 0$ ($\beta_0 := \hat{\beta}(0)$) and $\omega(1) = u(1,\beta_1) < 0$ ($\beta_1 := \hat{\beta}(1)$) so there must be an odd number of such zeroes, hence an odd number of such curve intersections.

For the nondegenerate case, we observe, from the theory of conic sections, that two such hyperbolae must coincide — impossible here — if they have more than two finite points in common, counting multiplicities. For the degenerate case (in which one or both of the curves becomes a straight line) one has the same conclusion. It follows (since 1 is the only odd number in $\{0,1,2\}$) that $\mathcal{C}_u$, $\mathcal{C}_v$ have exactly one intersection (α^*,β^*) within $[0,1]\times[0,1]$, necessarily within the interior $(0,1)\times(0,1)$. Since we have assumed that we are considering blending functions for which $\alpha(\cdot)$, $\beta(\cdot)$ are each strictly increasing when α, β, respectively, lies in $(0,1)$, it follows that there is a unique point (x^*,y^*) at which $\alpha(x^*) = \alpha^*$ and $\beta(x^*) = \beta^*$ — i.e., a unique stationary point for (5.18). We emphasize that this argument is quite specific to the sigmoid blending mechanism and to the case $m = 2$.

To see that this stationary point (x^*,y^*) is a global attractor, we note that we may apply the Poincaré–Bendixson theorem (as, with $m = 2$, we are in the plane with dynamics limited to a rectangle $\mathcal{B} = \overline{\alpha^{-1}(0,1)}\times\overline{\beta^{-1}(0,1)}$) which — as we already know that there is only one stationary point — informs us that any ω-limit set of the flow can only be (a) that stationary point, or (b) a periodic orbit, or (c) a homoclinic orbit. We wish to eliminate the latter two possibilities, each of which would give a trajectory bounding some region $\mathcal{R}$ in the interior of $\mathcal{B}$. To this end, we first note that

$$\nabla\cdot f := \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = u_\alpha \alpha' + v_\beta \beta' < 0$$

(strict inequality) since we have assumed $\alpha',\beta' > 0$ here and have, from (5.19) and (5.16), that

$$\begin{aligned} u_\alpha &= \beta\,(u_1 - u_2) + (1-\beta)\,(u_4 - u_3) < 0, \\ v_\beta &= \alpha\,(v_1 - v_4) + (1-\alpha)\,(v_2 - v_3) < 0. \end{aligned}$$

Next we note that since $\partial\mathcal{R}$ is a trajectory for (5.18) we have $d(x,y)$ parallel to f there so the normal $\mathbf{n}$ to $\partial\mathcal{R}$ is also orthogonal to f. Using Green's theorem, we would now have

$$0 = \oint_{\partial\mathcal{R}} f\cdot\mathbf{n}\,ds = \int\!\!\int_{\mathcal{R}} \nabla\cdot f\,dx\,dy < 0$$

which is a contradiction. Thus, $\{(x^*, y^*)\}$ is the only possible ω-limit set — i.e., is a global attractor for (5.18).

Since the flow always converges to this limit point (x^*, y^*), it follows — using (5.20) for the coefficients — that the limits in (4.4) necessarily exist since

$$c_1(s) = \alpha(x(s))\beta(y(s)) \to \alpha(x^*)\beta(y^*) =: \gamma_1^* \quad \text{as } s \to \infty,$$

etc. As in (5.15), this gives a sliding mode

(5.22)
$$f^* = f(x^*, y^*) = \sum_{k=1}^{4} \gamma_k^* w_k$$

with, of course, $(\gamma_{\boldsymbol{\sigma}}^*) := (\gamma_1^*, \gamma_2^*, \gamma_3^*, \gamma_4^*) \in \mathcal{C}_+$. Recall that we obtained (x^*, y^*) by solving $\alpha(x^*) = \alpha^*$ and $\beta(x^*) = \beta^*$ with (α^*, β^*) previously obtained by solving the quadratic system $u(\alpha, \beta) = 0 = v(\alpha, \beta)$ for $\alpha = \alpha^*$, $\beta = \beta^*$. Thus,

(5.23)
$$\gamma_1^* = \alpha^*\beta^*, \quad \gamma_2^* = (1-\alpha^*)\beta^*, \quad \gamma_3^* = (1-\alpha^*)(1-\beta^*), \quad \gamma_4^* = \alpha^*(1-\beta^*)$$

so we also have $(\gamma_{\boldsymbol{\sigma}}^*) = (\gamma_1^*, \gamma_2^*, \gamma_3^*, \gamma_4^*) \in \Gamma$. The uniqueness of the stationary point ensures that $\Gamma \bigcap \mathcal{C}_+$ is just the singleton $(\gamma_{\boldsymbol{\sigma}}^*)$.

As in connection with (4.5), we note that, having made this analysis on the fast time scale, with the vector fields 'frozen', we must restore the dependence on $z \in \mathcal{Z}$ to use this as a sliding mode on the natural time scale. We complete the proof by showing the regularity of f^*, as given by (5.22). First note that, once we have uniqueness, the pair (α^*, β^*) was obtained by solving $u(\alpha, \beta) = 0 = v(\alpha, \beta)$ so, by the Implicit Function Theorem, it depends smoothly (analytically!) on the field component values $[u_1, \ldots, v_4]$ appearing in (5.19) whence the dependence on $z \in \mathcal{Z}$ of this and of $(\gamma_{\boldsymbol{\sigma}}^*)$ as in (5.23) is as regular as theirs. For the regularity of $f^*(\cdot)$ we need only include consideration of the regularity of the parallel field components w_1, w_2 of f_1, f_2 which appear in (5.22). Essentially the same considerations apply to regularity of the sliding mode with respect to perturbation of the vector fields involved. ■

References

[1] E.A. Asarin and R.N. Izmailov, *Determining the sliding speed on a discontinuity surface,* Avtomatika i Telemekhanika no. 9 (1989), pp. 43–48. (transl., Automation and Remote Control **50** (1989), pp. 1181–1185.)

[2] A.F. Filippov, *Differential equations with the right-hand discontinuous on intersecting surfaces* (in Russian), Dif. Uravn. **15** (1979), pp. 1814–1823.

[3] A.F. Filippov, *Differential Equations with Discontinuous Righthand Sides,* Nauka, Moscow, 1985 [*transl.* Kluwer, Dordrecht, 1988].

[4] X. Ge, A. Nerode, W. Kohn, and J.B. Remmel, *Algorithms for chattering approximations to relaxed optimal controls,* preprint, 1995.

[5] T.I. Seidman, *Equations with hysteretically discontinuous right hand sides*, pp. 158–162 in *Models of Hysteresis,* (A. Visintin, ed.), Longman, Harlow (1993).

[6] T.I. Seidman, *Some limit problems for relays* in *Proc.* 1$^{\text{st}}$ *Int'l. Congress of Nonlinear Analysts*, (Tampa, 1992), to appear.

[7] V.I. Utkin, *Sliding Models and their Application in Variable Structure Systems*, Mir, Moscow, 1978.