

Chapter 6

Ideal and Non-ideal Sliding



Sliding—motion along the discontinuity threshold—is central to the most novel phenomena of nonsmooth dynamics. The theory was largely developed in Filipov’s work [51], but seems to originate in earlier Russian texts, perhaps first from G. N. Nikol’skii [109] (see also discussion in [4, 107, 148]). The standard definition would describe sliding as motion along an ideal threshold \mathcal{D} . We shall define it as follows.

Definition 6.1. A solution of an implementation (5.3) of a piecewise-smooth system (5.1) is said to **slide** if it evolves inside the layer \mathcal{D}^ε for a time $\Delta t = O(1)$ (i.e., a time not vanishing as $\varepsilon \rightarrow 0$).

This allows us to discuss sliding in implementations as well as in the ideal discontinuous system. Intuitively, sliding occurs because solutions tend towards some invariant set in the layer \mathcal{D}^ε around the discontinuity threshold \mathcal{D} . More precisely we can state the following.

Lemma 6.1. Consider system (5.1) on an open region W , defined piecewise on regions $\mathcal{R}_i = \{\mathbf{x} \in \mathbb{R}^n : i = \text{step}(\sigma(\mathbf{x})), \sigma \neq 0\}$ in terms of vector fields \mathbf{f}^i that are differentiable on $\sigma \geq -\varepsilon$ (for $i = 1$) and $\sigma \leq +\varepsilon$ (for $i = 0$), and a scalar function σ differentiable for all \mathbf{x} . Take an implementation of (5.1) on W according to Definition 5.1. If the discontinuity threshold $\mathcal{D} = \{\mathbf{x} : \sigma = 0\}$ is either attracting or repelling with respect to the vector fields \mathbf{f}^i for $\sigma \rightarrow 0$ on W , and the vector fields \mathbf{f}^i have non-vanishing components normal to \mathcal{D} for $\sigma \rightarrow 0$ on W , then there exists some $E > 0$ such that the layer \mathcal{D}^ε , where $\sigma = O(\varepsilon)$, is invariant inside W for $0 < \varepsilon < E$.

Proof. The attractivity of \mathcal{D} implies $\mathbf{f}^1(\mathbf{x}) \cdot \nabla\sigma < 0 < \mathbf{f}^0(\mathbf{x}) \cdot \nabla\sigma$, and repulsivity of \mathcal{D} implies the opposite signs, given that the normal components $\mathbf{f}^i \cdot \nabla\sigma$ are non-vanishing, evaluated at \mathbf{x} as $\sigma \rightarrow 0$. Since $\mathbf{f}^0(x)$ and $\mathbf{f}^1(x)$ are differentiable, at any point $\mathbf{x} \in W$ on $\sigma = 0$, there exists $\tilde{E}(\mathbf{x}) > 0$ such that $\mathbf{f}^1(\mathbf{u}) \cdot \nabla\sigma < 0 < \mathbf{f}^0(\mathbf{u}) \cdot \nabla\sigma$ for all \mathbf{u} such that $|\mathbf{u} \cdot \nabla\sigma| < \tilde{E}(\mathbf{x})$. Let $\tilde{E}(\mathbf{x})$ be the largest such value at each \mathbf{x} , and

let E be the infimum of all $\tilde{E}(\mathbf{x})$ for $\mathbf{x} \in \mathcal{D} \cap W$. Then for any $\varepsilon < E$ the vector fields satisfy $\mathbf{f}^1(\mathbf{x}) \cdot \nabla\sigma < 0$ for \mathbf{x} such that $\sigma < +\varepsilon$, and $\mathbf{f}^0(\mathbf{x}) \cdot \nabla\sigma > 0$ for \mathbf{x} such that $\sigma > -\varepsilon$, therefore the region $|\sigma(\mathbf{x})| < \varepsilon$ is invariant. \square

This suggests that if we can define a switching multiplier (some μ or ν), and a dynamics on it in the switching layer, then sliding constitutes an invariant of the dynamics confined to the layer, when the dynamics would otherwise carry trajectories through the layer. Even if we cannot define the layer dynamics in closed form, we can still consider the attracting or repelling objects it forms that constitute sliding.

Lemma 6.1 treats only the case when \mathcal{D} is attracting or repelling with respect to the flows outside it, but attractivity/repulsivity is neither necessary nor sufficient for sliding to take place. The extension of Lemma 6.1 to a discontinuity threshold \mathcal{D} formed by the intersection of manifolds $\mathcal{D}_1 \cap \dots \cap \mathcal{D}_m$, for example, is quite straightforward, this also has been mainly considered only under conditions of uniformly attraction of \mathcal{D} with respect to the surrounding flows, e.g., in [5, 38].

The situation in general is greatly more complicated. The discontinuity threshold \mathcal{D} need not be attracting or repelling for sliding to occur (as we saw in the example of ‘sticky genes’ in Sect. 1.3, see also [81]). Moreover, an intersection of manifolds $\mathcal{D}_1 \cap \dots \cap \mathcal{D}_m$ can be attracting without the individual thresholds \mathcal{D}_j being attracting, for example, if the flow spirals around a point $x_1 = x_2 = 0$ by crossing through discontinuity thresholds $x_1 = 0$ and $x_2 = 0$ (see, e.g., [36, 51]). The permutations are enormous and no substantial accounting of the possibilities has been made. Perhaps the most ambitious steps in this direction are in [66], where the authors classify the behaviours a solution can exhibit once it enters an intersection in a planar system.

The only general statement that can be made is that if the set \mathcal{G} from (5.2) has an intersection with the tangent space $\mathcal{T}_{\mathcal{D}}$ of the discontinuity threshold at a given point $\mathbf{x} \in \mathcal{D}$,

$$\mathcal{G}(\mathbf{x}) \cap \mathcal{T}_{\mathcal{D}}(\mathbf{x}) \neq \emptyset, \quad (6.1)$$

then sliding motion is possible in (5.1). Similarly if the set \mathcal{G}^ε of a given implementation from (5.3) has an intersection with the tangent space $\mathcal{T}_{\mathcal{D}}$ of the discontinuity threshold at a given point $\mathbf{x} \in \mathcal{D}$,

$$\mathcal{G}^\varepsilon(\mathbf{x}) \cap \mathcal{T}_{\mathcal{D}}(\mathbf{x}) \neq \emptyset, \quad (6.2)$$

then sliding motion is possible in the implementation of (5.1). A solution arriving at \mathcal{D}^ε evolves onto some attractor that lies in \mathcal{D}^ε , on which solutions $\mathbf{x}(t)$ evolve in a direction ε -close to the tangent space to \mathcal{D} . Thus we formalize the notion of sliding more precisely as follows.

Definition 6.2. A solution $\mathbf{x}(t)$ of (5.1) is said to **slide** along the discontinuity threshold \mathcal{D} if $\dot{\mathbf{x}}(t) \in \mathcal{G}(\mathbf{x}) \cap \mathcal{T}_{\mathcal{D}}(\mathbf{x})$. A solution $\mathbf{x}(t)$ of an implementation of (5.1) is said to **slide** along the discontinuity threshold \mathcal{D} (more strictly along the layer \mathcal{D}^ε approximating \mathcal{D}) if $\dot{\mathbf{x}}(t) \in \mathcal{G}^\varepsilon(\mathbf{x}) \cap \mathcal{T}_{\mathcal{D}}(\mathbf{x})$.

By saying $\dot{\mathbf{x}}(t) \in \mathcal{G}^\varepsilon(\mathbf{x}) \cap \mathcal{T}_{\mathcal{D}}(\mathbf{x})$, we mean that the tangent vector along a trajectory, $\dot{\mathbf{x}}(t)$, which must lie in $\mathcal{G}^\varepsilon(\mathbf{x})$, lies tangent to the discontinuity threshold \mathcal{D} , implying that $\mathbf{x}(t)$ evolves along \mathcal{D}^ε as t changes.

To find out whether sliding *will* occur in either situation requires a look at the dynamics, to find out whether the vector fields in the sets (6.1) and (6.2) possess attractors or repellers in the layer that can be followed by any solutions $\mathbf{x}(t)$. We explore the different approaches to this over Sects. 6.1 and 6.3.

The experiments in Chap. 4 show sliding under the implementations from Definition 4.1. In each case sliding occurs in a layer $\mathcal{D}_j^{\varepsilon_j}$ that forms an order ε_j -neighbourhood around an ideal discontinuity threshold $\mathcal{D}_j = \{\mathbf{x} : \sigma_j(\mathbf{x}) = 0\}$. Regardless of implementation, the attractivity of the layer implies the existence of local attractors inside it. The effects of Lemma 6.1 are seen in Figs. 4.1 to 7.4, with sliding occurring along the threshold $x = 0$ in Fig. 4.1, the thresholds $x_1 = 0 < x_2$ and $x_2 = 0 < x_1$ in Figs. 4.2 and 4.3, and the threshold $x_1 = 0 < x_2$ in Figs. 4.4, 4.5, and 4.6. With smoothing, these attractors are normally hyperbolic manifolds or equilibria. With hysteresis, the attractors are cycles oscillating between the boundaries $x_1 = \pm\varepsilon_1$ and $x_2 = \pm\varepsilon_2$, therefore reducible to return maps on those surfaces. For time stepping or delay the attractors can be bounded within definite ε_1 and ε_2 neighbourhoods of $x_1 = 0$ and $x_2 = 0$, respectively, and are described by piecewise-smooth two-dimensional maps on the plane.

Before returning to look at these experiments again closely, we need to build up a more general picture of the dynamics inside the switching layer that result from these definitions of layers, implementations, and sliding.

6.1 Sliding Perspective I: The Piecewise-Smooth System

Given the description of a piecewise-smooth system in terms of switching multipliers, as given by (5.2) with (5.8), we can analyse dynamics at the discontinuity as follows.

In (1.15) we defined the multipliers $\nu_j = \text{step}(\sigma_j)$ only as taking values $\nu_j \in [0, 1]$ at $\sigma_j = 0$. We shall now ask how each ν_j varies across the interval $[0, 1]$ and derive dynamics on them induced by the vector fields in $\sigma_j \neq 0$. (Clearly outside the discontinuity surface the ν_j s are piecewise-constant and so obey $\dot{\nu}_j = 0$ for any j).

On a discontinuity threshold \mathcal{D}_j , we can treat each ν_j as a ‘blow-up’ variable of the discontinuity set $\sigma_j = 0$. This is done by letting $\sigma_j = \varepsilon_j \nu_j$ for some small $\varepsilon_j \geq 0$, so that the discontinuity occurs across an interval $\sigma_j \in [0, \varepsilon_j] \rightarrow 0$ as $\varepsilon_j \rightarrow 0$. (This method is developed in [79, 81] but is essentially just a scaling of the quantity σ_j that maps its values on $\sigma_j \in [0, \varepsilon] \rightarrow 0$ to values $\nu_j \in [0, 1]$, and has no doubt been used earlier, e.g., in [148]. The term ‘blow-up’ itself appears to originate from singular perturbation literature [40]). We then use this to find the dynamics of ν_j on the interval $\nu_j \in [0, 1]$.

At a point where \mathcal{D} is a codimension m manifold, let $\mathbf{x}|_{\mathcal{D}} \in \mathbb{R}^{n-m}$ denote the space of \mathbf{x} restricted to \mathcal{D} . The multipliers ν_j in (5.10) lie on intervals $[0, 1]$, so the dynamics on \mathcal{D} can be said to take place inside a *switching layer*

$$\mathcal{D}^\varepsilon = \{ (\nu_1, \dots, \nu_m, \mathbf{x}|_{\mathcal{D}}) \in [0, 1]^m \times \mathbb{R}^{n-m} \}. \quad (6.3)$$

For economy of nomenclature we use the term ‘switching layer’ to describe both the parameterization of \mathcal{D} given by (6.3) in the piecewise-smooth system, and the region around \mathcal{D} given by (5.3) in the implementation, each with an associated small parameter ε . The concepts are closely related, and one may refer to the ‘switching layer of the piecewise-smooth system’ or the ‘switching layer of the implementation’ if necessary to avoid confusion.

At a point on a discontinuity threshold where $\sigma_1 = \dots = \sigma_m = 0$ for some $m \geq 1$, let us take local coordinates $\mathbf{x} = (X, \underline{x})$, where $X = (\sigma_1, \dots, \sigma_m)$ and $\underline{x} \in \mathbb{R}^{n-m}$, so $\mathbf{x}_{\mathcal{D}} = (0, \dots, 0, \underline{x})$. We then obtain on \mathcal{D} a switching layer

$$\mathcal{D}^\varepsilon = \{ (\nu_1, \dots, \nu_m, \underline{x}) \in [0, 1]^m \times \mathbb{R}^{n-m} \}, \quad (6.4)$$

with each multiplier ν_j constituting the blow-up variable of the set $\sigma_j = 0$, as some small parameter $\varepsilon_j \rightarrow 0^+$. The switching layer is n dimensional, and differentiating $\sigma_j = \varepsilon_j \nu_j$ according to $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; \nu_1, \dots, \nu_m)$ in these coordinates, given $\dot{\sigma}_j = \mathbf{F} \cdot \nabla \sigma_j = F_j$, we have

$$\begin{aligned} \varepsilon_j \dot{\nu}_j &= \dot{\sigma}_j = F_j(\varepsilon_1 \nu_1, \dots, \varepsilon_m \nu_m, \underline{x}; \nu_1, \dots, \nu_m) \\ &= F_j(0, \dots, 0, \underline{x}; \nu_1, \dots, \nu_m) + \mathcal{O}(\varepsilon_1, \dots, \varepsilon_m). \end{aligned} \quad (6.5)$$

Neglecting the higher order term for $\varepsilon_j \rightarrow 0$, we obtain a well-defined layer system on $X = (0, \dots, 0)$,

$$\varepsilon_j \dot{\nu}_j = F_j(0, \dots, 0, \underline{x}; \nu_1, \dots, \nu_m), \quad j = 1, \dots, m, \quad (6.6a)$$

$$\dot{\underline{x}} = \underline{F}(0, \dots, 0, \underline{x}; \nu_1, \dots, \nu_m), \quad (6.6b)$$

up to terms of order $\varepsilon_1, \dots, \varepsilon_m$, on the right-hand side. We recall that in this notation $\mathbf{F} = (F, \underline{F})$ and $F = (F_1, \dots, F_m) = (\mathbf{F} \cdot \nabla \sigma_1, \dots, \mathbf{F} \cdot \nabla \sigma_m)$.

In deriving this system we have fixed a very simple relationship between ν_j and σ_j , namely a linear (if singular as $\varepsilon_j \rightarrow 0$) mapping. We can do this without loss of generality because, through hidden terms, we are able to express any more complex functional relationship between some switching multiplier ν_j and the switching function σ_j using nonlinearity.

Say, for example, a vector field has a component $F_j = 1 + \text{step}(\sigma_j)$, representing perhaps the reaction force from an object stuck to a surface with $\sigma_j = 0$, and say that F_j is known to pass through zero twice as the function σ_j changes sign. Clearly the function $F_j = 1 + \nu_j$ does not satisfy this, as $\nu_j \in [0, 1]$ implies $1 + \nu_j \in [1, 2]$. The function $F_j = 1 + \nu_j - r\nu_j(1 - \nu_j)$, however, varies over $F_j \in [\frac{1}{4}(6 - r^{-1} - r), 2]$ for $r > 1$, and if, say, F_j is known to vanish at some $\nu_j = k$, then $r = (1 + k)/(1 - k)k$ provides this.

With the dynamics at a discontinuity threshold \mathcal{D} thus described by (6.6), sliding occurs if there exist fix points of (6.6a). These are sets of points satisfying $\dot{v}_j = 0$, and they generate *sliding manifolds*

$$\mathcal{M} = \left\{ (\underline{x}, v_1, \dots, v_m) \in \mathbb{R}^{n-m} \times [0, 1]^m \text{ such that } \begin{cases} F_j(0, \dots, 0, \underline{x}; v_1, \dots, v_m) = 0, & j = 1, \dots, m \end{cases} \right\}, \quad (6.7)$$

which are invariant wherever they are normally hyperbolic (see [81]). (Recall that $\mathbf{x}|_{\mathcal{D}} = (0, \dots, 0, \underline{x})$ denotes \mathbf{x} restricted to \mathcal{D} , and $\mathbf{F} \cdot \nabla \sigma_j = F_j$).

On \mathcal{M} the dynamics takes the form of a *sliding mode*, given by

$$0 = F_j(0, \dots, 0, \underline{x}; v_1, \dots, v_m), \quad j = 1, \dots, m, \quad (6.8a)$$

$$\dot{\underline{x}} = \underline{F}(0, \dots, 0, \underline{x}; v_1, \dots, v_m), \quad (6.8b)$$

Because (6.7) consists of m equations $F_1 = \dots = F_m = 0$, in m unknowns given by the switching multipliers v_1, \dots, v_m , they typically define a well-defined set \mathcal{M} inside a given switching layer. This set may consist of branches of different stability (determined by considering the eigenvalues of the matrix $\partial(F_1, \dots, F_m)/\partial(v_1, \dots, v_m)$), connected by non-hyperbolic points (where those eigenvalues have zero real part).

In a system of many switches $v_j = \text{step}(\sigma_j)$, a solution may evolve between places where the discontinuity threshold \mathcal{D} consists of an intersection of m different submanifolds \mathcal{D}_j , as defined in (5.11). On each different such region of \mathcal{D} for different m , we first blow up each \mathcal{D}_j into a switching layer, then derive the sliding modes, which occupy local sliding manifolds \mathcal{M} of dimension \mathbb{R}^{n-m} .

If there is nonlinear dependence on the multipliers v_j , then there may exist multiple equilibria, periodic or complex attractors, undergoing bifurcations and any other nonlinear phenomena inside the switching layer.

We will apply these ideas later when we look at some examples of linear versus nonlinear dynamics on the discontinuity threshold.

6.2 Sliding Perspective II: Hybrid Implementations

We define an implementation as *hybrid* if it cannot be expressed by means of a set of ordinary differential equations alone, but instead is given by a hybrid of the system

$$\dot{\mathbf{x}} = \mathbf{f}^i(\mathbf{x}) \quad \text{if} \quad \mathbf{x} \in \mathcal{R}_i^\varepsilon, \quad i \in Z_N, \quad (6.9)$$

along with a map

$$\dot{\mathbf{x}} = \mathbf{f}^i(\mathbf{x}) \quad \text{with} \quad i \mapsto \begin{cases} i & \text{if event}(\mathbf{x}; \varepsilon) = \text{false}, \\ \Psi(\mathbf{x}; i) & \text{if event}(\mathbf{x}; \varepsilon) = \text{true}, \end{cases} \quad (6.10)$$

with Z_N as before being some discrete set of N labels, and with the regions $\mathcal{R}_i^\varepsilon$ obeying $\mathcal{R}_i^\varepsilon \supset \mathcal{R}_i$, such that the switching layer \mathcal{D}^ε is formed by the overlap of two or more regions $\mathcal{R}_i^\varepsilon$, though $\mathcal{R}_i^0 = \mathcal{R}_i$.

The system evolves as $\dot{\mathbf{x}} = \mathbf{f}^i(\mathbf{x})$ until a condition ‘event($\mathbf{x}; \varepsilon$)’ is satisfied, then i is updated to a new mode $\Psi(\mathbf{x}; i)$. The hysteretic, delay, stochastic, and time-stepping implementations in Definition 4.1 are all of this type. Typically in such situations there is an implementation layer \mathcal{D}^ε on which the system may exist in more than one mode i , and its dynamics therefore depends not only on its state \mathbf{x} but also on the current mode, which therefore appears in the update map $\Psi(\mathbf{x}; i)$. In other implementations the dynamics in \mathcal{D}^ε is instead governed by a transition rule that only depends on \mathbf{x} , i.e., the state lies in a transition ‘mode’ and not in any mode i .

Let $\Omega \subset \mathcal{D}^\varepsilon$ be a set of points inside the switching layer on which ‘event($\mathbf{x}; \varepsilon$) = true’ is satisfied, and let $\mathbf{x}_1, \mathbf{x}_2, \dots$ denote a set of points inside Ω visited by $\mathbf{x}(t)$ at times t_1, t_2, \dots . Integrating between these provides a map

$$\mathbf{x}_n = \Phi(\mathbf{x}_{n-1}; \varepsilon) \quad \text{where} \quad \Phi : \Omega \mapsto \Omega. \quad (6.11)$$

We have been deliberately vague in defining the set Ω and map Φ , as the form of both depends on the implementation. In Chap. 7 we will see how these definitions apply to the ‘experiments’ from Chap. 4. Still, with these definitions we can derive some useful results, in particular without knowing the map (6.11) explicitly, we can derive some implications for the dynamics in the switching layer.

For sliding to occur, the map Φ must have an invariant set on Ω , but that set need not be unique, and by implication the sliding dynamics need not be unique.

According to the map (6.11), $\mathbf{x}(t)$ evolves in increments along each field \mathbf{f}^i , e.g.,

$$\begin{aligned} \mathbf{x}_n &= \mathbf{x}_{n-1} + \mathbf{f}^i(\mathbf{x}_{n-1})\Delta t + \mathcal{O}(\Delta t^2) \\ \text{if} \quad \mathbf{x}_{n-1} &\in \Omega_i(\mathbf{x}_{n-1}, \mathbf{x}_0, t_{n-1}). \end{aligned} \quad (6.12)$$

The mode i selected at each time increment can depend not only on the current state \mathbf{x}_{n-1} , but is also typically history dependent, depending on the initial state \mathbf{x}_0 and the current time t_{n-1} . The regions $\Omega_i \subset \Omega$ may therefore overlap. Because this selection has discontinuities wherever the mode i changes, this raises the possibility that attractors of the map can bifurcate in a far more arbitrary manner than in continuous or differentiable maps, able to make abrupt jumps in topology and periodicity (as in Fig. 2.2(ii), for example).

We can also derive the effect of an attractor on the system’s dynamics. Let $\mathbf{x}(t)$ evolve along an ε -infinitesimal neighbourhood of the discontinuity thresholds for a time interval $[0, T]$, switching between modes $i \in Z_N = \{1, 2, \dots, m\}$, at a sequence of times t_1, t_2, \dots, t_r , where $0 = t_0 < t_1 < t_2 < \dots < t_r = T$. Thus $\mathbf{x}(t)$ evolves along a different vector field \mathbf{f}^i in mode $i \in Z_N$ on each time interval $[t_{j-1}, t_j]$ for some $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, r\}$. Let μ_i denote the total proportion of the time T spent evolving along \mathbf{f}^i ,

$$\mu_i = \frac{1}{T} \sum_{n=1}^r \begin{cases} t_n - t_{n-1} & \text{if } \mathbf{x}_n \in \mathcal{R}_n, \\ 0 & \text{if } \mathbf{x}_n \notin \mathcal{R}_n. \end{cases} \quad (6.13)$$

Let $\gamma_i(\mathbf{x}_i) = 1$ if \mathbf{x}_n is currently in mode \mathbf{f}^i and $\gamma_i(\mathbf{x}_i) = 0$ otherwise. Then the total change in $\mathbf{x}(t)$ over the time increment $T = \sum_{n=1}^r (t_n - t_{n-1})$ is

$$\begin{aligned} \frac{\Delta \mathbf{x}}{T} &= \frac{1}{T} \sum_{n=1}^r \Delta \mathbf{x}_{n-1} \\ &= \frac{1}{T} \sum_{n=1}^r \gamma_i(\mathbf{x}_{n-1}) \mathbf{f}^i(\mathbf{x}_{n-1}) (t_n - t_{n-1}) + \mathcal{O}(\Delta t^2) \\ &= \sum_{i=1}^r \mu_i \mathbf{f}^i(\mathbf{x}) + \mathcal{O}(T), \end{aligned} \quad (6.14)$$

where $\mu_i \geq 0$ and $\sum_{i=1}^N \mu_i = 1$. In the limit $T \rightarrow 0$ this gives an effective equation of motion,

$$\dot{\mathbf{x}} = \mathbf{F}^{\text{co}}(\mathbf{x}; \mu_1, \dots, \mu_N) := \sum_{i=1}^N \mu_i \mathbf{f}^i(\mathbf{x}). \quad (6.15)$$

Comparing to (5.5a), we see that as the coefficients μ_i vary over $[0, 1]$ the effective vector field \mathbf{F}^{co} traces out the convex hull $\mathcal{F}(\mathbf{x})$ of the \mathbf{f}^i 's.

Using these effective equations of motion we can understand a basic separation of scales that distinguishes motion *across* and *along* the discontinuity threshold. Assume that \mathbf{x} evolves along an ε -neighbourhood of the intersection of discontinuity thresholds $\sigma_1 = \dots = \sigma_m = 0$. Define scaled coordinates $\mathbf{x} = (u, v, \dots, w, \underline{x})$, where $u = \sigma_1/\varepsilon$, $v = y/\varepsilon$, \dots , $w = \sigma_m/\varepsilon$, and $\mathbf{x} \in \mathbb{R}^{n-m}$. Let $\mathbf{F}^{\text{co}} = (F_u^{\text{co}}, \dots, F_w^{\text{co}}, \underline{F}^{\text{co}})$, then $\dot{\mathbf{x}} = \mathbf{F}^{\text{co}}$ becomes

$$\begin{aligned} \varepsilon \dot{u} &= F_u^{\text{co}}(\varepsilon u, \dots, \varepsilon w, \underline{x}; \mu_1, \dots, \mu_m) \\ &= F_u^{\text{co}}(0, \dots, 0, \mathbf{x}; \mu_1, \dots, \mu_m) + \mathcal{O}(\varepsilon), \\ &\vdots \\ \varepsilon \dot{w} &= F_w^{\text{co}}(\varepsilon u, \dots, \varepsilon w, \underline{x}; \mu_1, \dots, \mu_m) \\ &= F_w^{\text{co}}(0, \dots, 0, \mathbf{x}; \mu_1, \dots, \mu_m) + \mathcal{O}(\varepsilon), \end{aligned} \quad (6.16a)$$

$$\begin{aligned} \dot{\underline{x}} &= \underline{F}^{\text{co}}(\varepsilon u, \dots, \varepsilon w, \underline{x}; \mu_1, \dots, \mu_m) \\ &= \underline{F}^{\text{co}}(0, \dots, 0, \mathbf{x}; \mu_1, \dots, \mu_m) + \mathcal{O}(\varepsilon). \end{aligned} \quad (6.16b)$$

(Here part (a) labels the fast equations $\varepsilon[\dot{\quad}] = \dots$ and (b) the slow equation $\dot{\underline{x}} = \dots$). The (u, \dots, w) coordinates therefore evolve on a fast timescale $\tau = t/\varepsilon$. Denoting the derivative with respect to τ by a prime, the system instead becomes

$$\begin{aligned} u' &= F_u^{\text{co}}(0, \dots, 0, \mathbf{x}; \mu_1, \dots, \mu_m) + \mathcal{O}(\varepsilon) , \\ &\vdots \\ w' &= F_w^{\text{co}}(0, \dots, 0, \mathbf{x}; \mu_1, \dots, \mu_m) + \mathcal{O}(\varepsilon) , \end{aligned} \quad (6.17a)$$

$$\underline{x}' = \mathcal{O}(\varepsilon) . \quad (6.17b)$$

When we simulate the system on the τ -timescale, the variables (u, \dots, w) evolve across the $\mathcal{O}(\varepsilon)$ space of the switching layer, while the variables \mathbf{x} remain quasi-static.

To identify the coefficients μ_i in (6.14), we therefore simulate (6.17) for a time interval T , keeping \underline{x} fixed. If the time interval T can be taken long enough that the coefficients μ_i as calculated by (6.13) reach a steady state, their values define an effective equation of motion (6.15) at any given \mathbf{x} . If one of the μ_i takes a value of unity, with all $\mu_{j \neq i} = 0$, then the system is determined to have *crossed* the discontinuity threshold. If one or more μ_i settle to steady values between 0 and 1, then the simulation of the fast (u, \dots, w) subsystem of (6.17) must have reach a steady state or other attractor inside the switching layer, and is said to be *sliding* along the discontinuity.

The maps (6.11) and their invariants may not in general have any closed form expression that can be determined from the system (5.1), but must be discovered by simulation and approximated.

Although (6.16) is similar formally to (6.6), the latter describes a continuous flow on $\sigma_1 = \dots = \sigma_m = 0$, while the former describes a hybrid implementation that jumps between the 2^m different modes specified by (6.16) when each μ_j takes a value 0 or 1 in the neighbourhood of $\sigma_1 = \dots = \sigma_m = 0$.

6.3 Sliding Perspective III: Smoothed Implementations

If switching is implemented by a smooth process, then we can proceed by steps that are actually very similar to Sect. 6.1. This has the advantage that the analysis then follows standard methods of singular perturbation theory, but this familiarity disguises unobvious ambiguities that accompany smoothing. Though we can describe these to some extent here, we are still learning what kind of dynamics, and more specifically what kind of singularities, persists under smoothing.

We begin with the system $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; v_1, \dots, v_m)$ in terms of switching multipliers $v_j = \text{step}(\sigma_j)$. To smooth the discontinuity we simply make the replacement $v_j \mapsto \phi^{\varepsilon_j}(\sigma_j)$, where $\phi^{\varepsilon_j}(\sigma_j)$ is a smooth monotonic function, satisfying $\phi^{\varepsilon_j}(\sigma_j) = \text{step}(\sigma_j) + \mathcal{O}(\varepsilon_j)$ for $|\sigma_j| \geq \varepsilon_j$.

Observe that for $|\sigma_j| \geq \varepsilon_j$ this definition gives $\phi^{\varepsilon_j}(\sigma_j) = \phi^1(\sigma_j/\varepsilon_j)$, so let assume this also holds for $|\sigma_j| < \varepsilon_j$. The dynamics of the quantity v_j is found simply by differentiating and applying the chain rule, $\dot{v}_j = \varepsilon_j^{-1}(\partial\phi^{\varepsilon_j}/\partial\sigma_j)\dot{\sigma}_j = \hat{\varepsilon}_j^{-1}\dot{\mathbf{x}} \cdot \nabla\sigma_j = \hat{\varepsilon}_j^{-1}\mathbf{F} \cdot \nabla\sigma_j = \hat{\varepsilon}_j F_j$, where $\hat{\varepsilon}_j = \varepsilon_j(\partial\phi^{\varepsilon_j}/\partial\sigma_j)^{-1}$.

The result is formally that in (6.5)–(6.6), except that the definition of v_j now differs and, more crucially, the small quantity $\hat{\varepsilon}_j$ is now a function of σ_j . We have

$$\hat{\varepsilon}_j(\sigma_j)\dot{v}_j = F_j(0, \dots, 0, \underline{x}; v_1, \dots, v_m) + \mathcal{O}(\varepsilon_1, \dots, \varepsilon_m) , \quad (6.18a)$$

$$\dot{\underline{x}} = \underline{F}(0, \dots, 0, \underline{x}; v_1, \dots, v_m) + \mathcal{O}(\varepsilon_1, \dots, \varepsilon_m) . \quad (6.18b)$$

By the definition of ϕ^{ε_j} the quantity $\hat{\varepsilon}_j$ is non-zero on the layer, where $|\sigma_j| < \varepsilon_j$. We can refine this if we limit the derivative of ϕ^{ε_j} away from zero, for example, choose ϕ^{ε_j} such that $\partial\phi^{\varepsilon_j}/\partial\sigma_j > K$ for $|\sigma_j| < 1 - \varepsilon_j^p$, for fixed $K > 0$ and $p \geq 1$ such that $\varepsilon_j/K \rightarrow 0$ as $\varepsilon_j \rightarrow 0$. Then $\hat{\varepsilon}_j$ behaves as a small quantity $\hat{\varepsilon}_j/K = \mathcal{O}(\varepsilon_j)$ for $\sigma_j \in [-1 + \varepsilon_j^p, +1 - \varepsilon_j^p]$.

We see that this is analogous to the system (6.6) obtained by piecewise-smooth methods, and in fact they can be shown to be equivalent in the limit $\varepsilon_1, \varepsilon_2 \rightarrow 0$, see [81].

Equations of the form (6.18) can be found in singular perturbation studies of climate and gene regulation, e.g., [96, 103]. An equivalent form in common use when using Sotomayor–Teixeira regularization [136] is to define a parameter $u_j = \sigma_j/\varepsilon_j$, to obtain instead

$$\varepsilon_j \dot{u}_j = F_j(0, \dots, 0, \underline{x}; v_1, \dots, v_m) + \mathcal{O}(\varepsilon_1, \dots, \varepsilon_m) , \quad (6.19a)$$

$$\dot{\underline{x}} = \underline{F}(0, \dots, 0, \underline{x}; v_1, \dots, v_m) + \mathcal{O}(\varepsilon_1, \dots, \varepsilon_m) . \quad (6.19b)$$

In either case (6.18) or (6.19), analysis proceeds using standard concepts from geometric singular perturbation theory, see, e.g., [49, 86]. If we assume all of the ε_j s are of the same order, that is every ratio $\varepsilon_i/\varepsilon_j$ is non-vanishing as $\varepsilon_i, \varepsilon_j \rightarrow 0$ for $i, j = 1, \dots, m$, then the analysis is closely analogous to that of the piecewise-smooth system in Sect. 6.1. The switching layer of the implementation is given as in (6.4) treating the v_j as variables, or the same expression with v_j replaced by u_j in the alternative variables. The slow-fast system has a critical manifold, corresponding precisely to the sliding manifold \mathcal{M} in (6.7), where the fast v_j subsystem (6.6a) vanishes. According to Fenichel’s theory [49], wherever \mathcal{M} is normally hyperbolic with respect to the fast v_j subsystem, for $\varepsilon_j > 0$, there exists an invariant manifold $\mathcal{M}_{\varepsilon_j}$ in an ε_j -neighbourhood of \mathcal{M} . The dynamics on \mathcal{M} is precisely the sliding dynamics (6.8), and moreover the dynamics on $\mathcal{M}^{\varepsilon_j}$ is topologically equivalent to (6.8).

If all of the ε_j s are of different orders, then the dynamics in the switching layer will be more intricate, involving a separation onto more timescales, but still falls under standard methods of singular perturbation theory. The case $\varepsilon_j = \varepsilon_1^j$ for $j = 1, \dots, m$, for instance, falls under Fenichel’s analysis in [49]. The author is not aware of any studies to date applying such many timescale dynamics to piecewise-smooth problems.

The Sotomayor–Teixeira approach of replacing the switching multipliers v_j by smooth (but non-analytic) functions $\phi^{\varepsilon_j}(\sigma_j)$ can be weakened so that the functions $\phi^{\varepsilon_j}(\sigma_j)$ are analytic. It is then impossible for these functions to be constant outside the switching layer, so they require defining to $\phi^{\varepsilon_j}(\sigma_j) = \text{step}(\sigma_j) + E$ where E is

small, for example, $E = O(\varepsilon_j)$ or $O(\varepsilon_j/\sigma_j)$ or $O(e^{-\sigma_j/\varepsilon_j})$. This is often the case in applications. One example is in [96], where $\phi^{\varepsilon_j}(\sigma_j) = \frac{1}{2} + \frac{1}{2} \arctan(\sigma_j/\varepsilon) = \text{step}(\sigma_j) + O(\varepsilon_j/\sigma_j)$. Another example is in [103, 117], where $\phi^{\varepsilon_j}(\sigma_j) = Z(\sigma_j + 1) = \text{step}(\sigma_j) + O(\sigma_j^{-1/\varepsilon_j})$ in terms of the Hill function $Z(w) = 1/(1 + w^{-1/\varepsilon_j})$ [70]. Both of these functions are analytic and are asymptotic to $\text{step}(\sigma_j)$ for large argument.

Good examples of these methods applied to genetic models like that of Sect. 4.2 can be found in [41, 76, 102, 103, 117]. They tell a story similar to that obtained by piecewise-smooth analysis, but an appreciation of possible hidden terms would bring more insight into the robustness of these studies. Hill functions as a class are sometimes used without rigorous justification from the biology, and in such cases the possible differences between alternate sigmoid functions can be calculated as hidden terms, including representing the different between Hill functions of different stiffnesses (different powers $1/\varepsilon_j$).

Hidden terms survive when we smooth a discontinuity, and have the interpretation that they vanish (asymptotically at least) outside the discontinuity threshold. If we smooth by replacing $v_j \mapsto \phi^{\varepsilon_j}(\sigma_j) = \text{step}(\sigma_j) + E(\varepsilon_j)$, then the term $v(v - 1)$ which from (5.16) typically characterizes hidden terms, simplifies to

$$v(v - 1) \mapsto 2 \text{sign}(\sigma_j)E(\varepsilon_j) + e^2(\varepsilon_j) .$$

Hence the hidden term is of order $E(\varepsilon_j)$, vanishing (asymptotically) outside the discontinuity threshold with E .

Hidden terms therefore allow us to distinguish between different kinds or rates of switching according to different methods of smoothing. For instance, consider the one-dimensional system

$$\dot{x} = F(x; v_{(r)}) = a(x) + vb(x) , \quad (6.20)$$

defined in terms of a switching multiplier

$$v = v_{(r)} := \lim_{\varepsilon \rightarrow 0} \phi_{(r)}^\varepsilon(x) , \quad (6.21)$$

for different smooth functions $\phi_{(1)}^\varepsilon(x)$, $\phi_{(2)}^\varepsilon(x)$, \dots , such that $v_{(r)} \rightarrow \text{step}(x)$ for any r . Does it matter how we choose the function $\phi_{(r)}^\varepsilon$, or do we always obtain the same piecewise-smooth system (6.20) in the limit $\varepsilon \rightarrow 0$?

As examples consider the following sigmoid quantities

$$\phi_{(0)}^\varepsilon(x) = \frac{1}{2} + \frac{x/\varepsilon}{2\sqrt{1+(x/\varepsilon)^2}} \quad (6.22a)$$

$$\phi_{(1)}^\varepsilon(x) = \frac{1}{2} + \frac{x/\varepsilon}{2\sqrt{1+(x/\varepsilon)^2}} + \frac{A(x)}{2(1+(x/\varepsilon)^2)^k} , \quad (6.22b)$$

$$\phi_{(2)}^\varepsilon(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(x/\varepsilon) , \quad (6.22c)$$

$$\phi_{(3)}^\varepsilon(x) = \frac{1}{2} + \frac{1}{2} \tanh(x/\varepsilon) , \quad (6.22d)$$

where $k > 0$ and $A(x)$ is a smooth function of x .

We will show the following.

Lemma 6.2. *We can write each system $\dot{x} = F(x; \nu_{(r)})$ from (6.20) as*

$$\dot{x} = F(x; \nu_{(r)}) = F(x; \text{step}(x)) + H_{(r)}^\varepsilon(x), \quad (6.23)$$

an asymptotic expansion whose tail satisfies $H_{(r)}^\varepsilon(x) \rightarrow 0$ as $\varepsilon \rightarrow 0$ for $x \neq 0$. This can be re-written in an ε -independent form as

$$\dot{x} = F(x; \nu_{(r)}) = F(x; \nu_{(0)}) + H_{(r)}(x; \nu_{(0)}), \quad (6.24)$$

with hidden terms satisfying $H_{(r)}(x; 0) = H_{(r)}(x; 1) = 0$, but with $H_{(r)}(x; \nu)$ begin non-vanishing in the layer $|x| < \varepsilon$ for $r = 1, 2, 3$.

Proof. The proof is directly by asymptotic expansion and straightforward calculations, so for brevity we place some of the details in Appendix C. The expansions of the sigmoid functions $\phi_{(r)}^\varepsilon$ for large argument all take the form $\phi_{(r)}^\varepsilon(x) = \text{step}(x) + \mathcal{O}(\varepsilon/x)$ (the precise expressions are given in Appendix C, and in fact for $r = 3$ the error is $\mathcal{O}(e^{-\varepsilon/x})$). Substituting these into (6.20) gives (6.23) with the tail of the expansion given by $H_{(r)}^\varepsilon(x) = \mathcal{O}(\varepsilon/x)$.

This means that the systems (6.20) with (6.21) and (6.22) are equivalent for $x \neq 0$. For $x = 0$, however, these asymptotic series diverge. To compare the different systems at $x = 0$, the second part of the theorem instead seeks a form of the expansions that is independent of ε , by expressing them all in terms of $\nu_{(0)}$.

To do this we first rearrange (6.22a) to find that

$$x/\varepsilon = (\phi_{(0)}^\varepsilon - \frac{1}{2}) / \sqrt{\phi_{(0)}^\varepsilon(1 - \phi_{(0)}^\varepsilon)},$$

followed by substituting this into the expressions in (6.22) and then into (6.20). To obtain (6.24) for $r = 1$ is then just a matter of algebra. To obtain (6.24) for $r = 2, 3$, it is better to substitute into the asymptotic expansions for each $\phi_{(r)}^\varepsilon$, thus expressing them in terms of $\phi_{(0)}^\varepsilon$. We then replace $\phi_{(0)}^\varepsilon$ with $\nu_{(0)}$. The algebra is set out in Appendix C, giving (6.24) for $r = 1, 2, 3$, with

$$H_{(1)}(x; \nu) = (4h)^k A(x)b(x), \quad (6.25a)$$

$$H_{(2)}(x; \nu) = \left\{ h C_1(2\nu - 1) + \sqrt{h} C_2(2\nu - 1) \right\} b(x), \quad (6.25b)$$

$$H_{(3)}(x; \nu) = \left\{ h C_1(2\nu - 1) + e^{-|2\nu-1|/\sqrt{h}} C_3(2\nu - 1) \right\} b(x), \quad (6.25c)$$

where $h = \nu(1 - \nu)$,

in terms of functions $C_i(2\nu - 1)$ that are finite valued for all $\nu \in [0, 1]$, given in Appendix C. Clearly $H_{(r)}(x; 0) = H_{(r)}(x; 1) = 0$ in each case. These expressions are now independent of ε and therefore remain well-defined in terms of $\nu_{(0)}$ as $\varepsilon \rightarrow 0$. \square

Compare the hidden terms in (6.25) with the expression (5.12) in (5.3). Note how the term “ $\nu(1 - \nu)$ ” appears throughout the hidden terms $H_{(r)}(x; \nu)$ in (6.25), but also

demonstrates more general forms that hidden terms can take than we derived by polynomial expansion in (5.16).

This merely demonstrates that a piecewise-defined function corresponds not to one unique function of a switching multiplier ν or limiting smooth function $\phi^\varepsilon(x)$, but a whole class of such functions. Comparing to (5.3) we see that the difference between the alternate smoothings $\mathbf{F}(\mathbf{x}; \nu_{(r)})$ lies in hidden terms.

Our interest, of course, concerns the dynamical implications of these hidden terms, left behind from the asymptotic approximations above. It should be quite clear that they can affect the system's dynamics. There are examples in [78, 79, 81, 83] of hidden terms deciding whether solutions slide along or cross through a discontinuity threshold. The anomalous sliding we described in Sects. 1.2 and 1.3 came from hidden terms, and more generally that can take all manner of non-trivial forms. An interesting example is given by taking (6.22b) and letting $A(x)$ be a matrix. The following example shows how this can destabilize an equilibrium under smoothing. Consider the piecewise-linear problem

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} -3 \\ ay + z \\ az - y \end{pmatrix}, \quad \nu = \text{step}(x), \quad (6.26)$$

for $a < 0$, which has sliding modes satisfying $\nu = 1/3$, with an attracting focus equilibrium at $y = z = 0$. If we smooth this system by replacing ν with $\phi_{(0)}^\varepsilon$, then we obtain a topologically equivalent system, with an attracting focus equilibrium on an invariant manifold, where $\phi_{(0)}^\varepsilon = 1/3$. Consider instead smoothing by replacing ν with $\phi_{(1)}^\varepsilon$, and let

$$A = c \begin{pmatrix} -1/3 & 0 & 0 \\ 0 & a & -1 \\ 0 & 1 & a \end{pmatrix}, \quad (6.27)$$

for small $c > 0$. As this is now a smooth system it succumbs to standard stability analysis. The system has an equilibrium at $(x, y, z) = (x_*, 0, 0)$, where $\phi_{(0)}(x_*/\varepsilon) = \frac{1}{3} + \frac{8c}{27} + \mathcal{O}(c^2)$. This has eigenvalues $-\frac{3}{\varepsilon}(1 - \frac{4c}{9})\phi'_{(0)}(x_*/\varepsilon)$ and $\frac{1}{3}(a \pm i) + \frac{8c}{27}(3 + a + 3a^2 \pm i)$ to order c^2 . This implies that for

$$\frac{-3a}{8[(a+1)^2 - \frac{5}{3}a]} < c < \frac{9}{4},$$

the equilibrium will de-stabilize in the (y, z) directions, becoming a saddle-focus as depicted in Fig. 6.1.

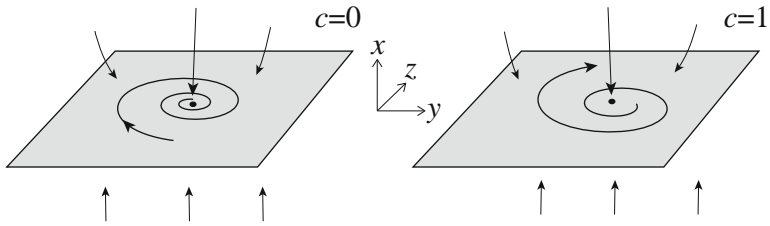


Fig. 6.1 A focus destabilized by smoothing

In this section we have seen some of the less obvious complexity of switching and sliding dynamics when considered from different viewpoints, expressed through layers, nonlinearity, and implementations. Let us now return to see what insight these give us into the ambiguities of the examples in Chap. 4.