

Feedback modal control of partial differential equations

Thomas I. Seidman

Abstract. For hybrid systems in which control consists of selection from a discrete finite set of *modes*, a somewhat unfamiliar formulation is needed for analysis of the possibility of closed loop (feedback) control. We are here concerned to examine the desiderata for such feedback from the viewpoint of descriptive modeling of implementation in a PDE context. A principal result is global existence, in an appropriate sense, for the implemented closed loop control system. A problem of transport on a graph is then presented to show how the relevant hypotheses might be satisfied in a PDE example.

Mathematics Subject Classification (2000). 93A30, 93B12, 47J40, 70K70.

Keywords. modeling, multiscale, hybrid systems, switching, modes, discontinuities, differential equations, feedback, Zeno phenomena.

1. Introduction

Consider a collection of partial differential equations which we take, somewhat arbitrarily, to have the form

$$\dot{x} = \mathbf{A}_j x + f_j(x) \quad (j \in \mathcal{J}) \quad (1.1)$$

where each \mathbf{A}_j is a suitable differential operator. Now imagine a system whose evolution is governed, over interswitching intervals, by one or another of these; we call this a *hybrid system* and take the discrete *modal index* $\mathbf{j} = \mathbf{j}(t)$ to be a component of the system state along with the ‘continuous component’ x . We will be considering the modal transitions $[\mathbf{j}(t-) = j] \curvearrowright [\mathbf{j}(t+) = j']$ as our control mechanism for the system. In particular, we will be concerned with the possibility of closed loop operation of such a control system.

Such hybrid systems are already a much-studied area of interest in the lumped parameter (ordinary differential equation) context, although much of the underlying theory remains open. Although one of the early analyses of such systems [7] was, indeed, motivated by a PDE example (not in a control context), we note that

very little has been done so far for the analysis of switching control for distributed parameter problems governed by partial differential equations, despite the fact that significant applications come easily to mind, involving the use of such familiar ON/OFF control devices as valves and pumps, light switches and thermostats, traffic signals, etc.

[One might also consider systems with $\mathbf{j} = \mathbf{j}(t, s)$ pointwise, compare [9] where this becomes a free boundary problem for the set $\{(t, s) : \mathbf{j}(t, s) = j\}$. Here we will restrict our attention to situations where the switching may be viewed as global with a finite number of modes — as is the case, e.g., for traffic signals if we take each mode as specifying the configuration of signal states for the entire road network. As with hybrid ODEs, the index \mathbf{j} will be a function of t alone.] Many, but not all, of the relevant aspects of the analysis are then independent of dimensionality.

In the context of feedback for PDEs, the regularity of the sensor inputs being considered may be significant even to know that solutions exist. Our formulation reflects a concern for the modeling of such systems. This will be very much a question of time scales: we are assuming that the switching itself takes place on a time scale more rapid than our modeling concerns but that the interswitching intervals are on the scale of interest.

While other considerations may also be of interest — e.g., controllability or stabilization to a small region — we here envision three canonical results:

Theorem 1: *Under appropriate hypotheses, treating $\mathbf{j}(\cdot)$ as data, the system will be well-posed in some suitable sense.*

Again treating open-loop control with a suitable cost functional,

Theorem 2: *Under appropriate hypotheses there exists an optimal control \mathbf{j} , minimizing the cost. For the autonomous infinite horizon problem this can be obtained by a kind of feedback.*

Modifying that notion of *feedback* to be based on suitable sensors,

Theorem 3: *Under appropriate hypotheses the feedback controlled system will be well-posed in some suitable sense.*

Much of the paper is devoted to explaining what these should mean.

2. The formulation

We begin by noting an important distinction between *descriptive* and *prescriptive* modes of modeling: the first is what a scientist does in trying to understand the various patterns arising in the world; the second is what a composer or an engineer does in designing (artificial) patterns for various purposes. E.g., in viewing leonine behavior the first is the modality of the naturalist while the second is the approach of a lion tamer. In this section we provide a formal prescriptive model for lion tamers, while aware that a naturalist's comments will later be complementary in describing how lions can behave.

The formal elements of the underlying system consist of

$$\begin{aligned}
 \text{(E)} \quad & \begin{aligned}
 & \text{a. a finite index set } \mathcal{J} \text{ and a corresponding set of state spaces } \mathcal{X}_j. \\
 & \text{b. a set of information spaces } \mathcal{Y}_j \text{ with sensor maps } Y_j : \mathcal{X}_j \rightarrow \mathcal{Y}_j. \\
 & \text{c. a nonempty action set } \mathcal{A}_j(y) \subset \mathcal{J} \text{ for each } j \in \mathcal{J}, y \in \mathcal{Y}_j. \\
 & \text{d. a continuous transition map } \mathbf{f} : [j, x] \mapsto [j', x'] \text{ with } x' \in \mathcal{X}_{j'} \\
 & \text{defined when } j \in \mathcal{J}, x \in \mathcal{X}_j \text{ and } j' \in \mathcal{A}_j(Y_j(x)). \\
 & \text{e. a set of dynamical systems = modes } \pi_j, \text{ each satisfying the} \\
 & \text{causality condition} \\
 & \pi_j(t, s, \xi) = \pi_j(t, r, \pi_j(r, s, \xi)) \quad \text{for } t \geq r \geq s, \quad \xi \in \mathcal{X}_j \quad (2.1)
 \end{aligned}
 \end{aligned}$$

We assume throughout that each $\mathcal{X}_j, \mathcal{Y}_j$ is a complete metric space and that each π_j and \mathbf{f} is continuous; at this point we impose no continuity requirements on Y_j . [We expect the modes of (E)-e. to be given as in (1.1) so continuity of π_j just means well-posedness. Note, however, that any relevant boundary data is then to be included in specification of the mode.]

We may anthropomorphize the feedback as a *controller* who knows the current mode and sensor values $j = \mathbf{j}(t)$ and $y = \mathbf{y}(t) = Y_j(\mathbf{x}(t))$ and, based on this, continually selects the mode. Of course the controller's choices at any moment are restricted to the available *control actions*: to remain in the current mode j or to make a transition $j \leadsto j'$ on the fast scale; the *switching rules* are just that this selection always be taken from the action set $\mathcal{A}_j(y)$. It will also be convenient to introduce the sets

$$\mathcal{S}^j := \{y \in \mathcal{Y}_j : j \in \mathcal{A}_j(y)\}, \quad \mathcal{C}^{j \leadsto j'} := \{y \in \mathcal{Y}_j : j' \in \mathcal{A}_j(y)\}, \quad (2.2)$$

noting that the required nonemptiness of each $\mathcal{A}_j(y)$ ensures that

$$\mathcal{S}^j \cup \mathcal{B}^j = \mathcal{Y}_j \quad \text{where } \mathcal{B}^j = \bigcup_{j' \neq j} \mathcal{C}^{j \leadsto j'}. \quad (2.3)$$

We refer to the specification of the sets $\{\mathcal{S}^j, \mathcal{B}^{j \leadsto j'}\}$ as the *switching diagram* for mode j and to these collectively ($j \in \mathcal{J}$) as the *controlling feedback diagram* for the system.

Intuitively, the operation of such a feedback controlled system should produce finitely many switching times $\{t_\nu : \nu = 1, \dots, N\}$ in any time interval $[0, T]$ with a modal transition $j_\nu \leadsto j_{\nu+1}$ at each t_ν — thus partitioning $[0, T]$ into the *interswitching intervals* $\mathcal{I}_\nu = [t_{\nu-1}, t_\nu]$ with $0 = t_0 \leq t_1 \leq \dots$. A *solution* of such a feedback system on the time interval $[0, T]$ would then be a triple of functions $[\mathbf{j}(\cdot), \mathbf{x}(\cdot), \mathbf{y}(\cdot)]$ such that

- (S)
- a. $\mathbf{j}(\cdot)$ is piecewise constant with $\mathbf{j}(t) = j_\nu \in \mathcal{J}$ for t in each interswitching interval $\mathcal{I}_\nu = [t_{\nu-1}, t_\nu]$; at each t one will have $\mathbf{x}(t) \in \mathcal{X}_{\mathbf{j}(t)}$ and $\mathbf{y}(t) = Y_{\mathbf{j}(t)}(\mathbf{x}(t)) \in \mathcal{Y}_{\mathbf{j}(t)}$.
 - b. the switching times are discrete: finitely many in any $[0, T]$.
 - c. switching $[j_{\nu-1} \curvearrowright j_\nu$ at $t_\nu]$ occurs only if $j_\nu \in \mathcal{A}_{j_{\nu-1}}(\mathbf{y}(t_{\nu-}))$ while for t in the interior of \mathcal{I}_ν one must have $\mathbf{y}(t) \in \mathcal{S}^{j_{\nu-1}}$.
 - d. at each switching time t_ν

$$\mathbf{x}(t_\nu+) = \mathbf{f}(\mathbf{x}(t_\nu-); j_{\nu-1} \curvearrowright j_\nu). \quad (2.4)$$
 - e. on each interswitching interval \mathcal{I}_ν we have $\mathbf{x}(t) \in \mathcal{X}_{j_{\nu-1}}$ with
$$\mathbf{x}(t) = \pi_{j_{\nu-1}}(t - t_{\nu-1}, \mathbf{x}(t_{\nu-1}+)). \quad (2.5)$$

Deferring further discussion to the next section, note that (S)-a. will admit the possibility of degenerate interswitching intervals ($t_\nu = t_{\nu-1}$), for which we formally take $\mathbf{x}(t_\nu-) = \mathbf{x}(t_{\nu-1}+)$ with no evolution. The occurrence of infinitely many transitions within a finite period is known as a *Zeno phenomenon* and this possibility would be a major technical difficulty for the theory; (S)-b. requires that this does not occur in the problems we consider.

Remark 2.1. We note a few generalizations which can be included within the framework of (E), (S).

We have formulated the feedback to depend only on the current sensor values $Y_j(t) \in \mathcal{Y}_j$, without memory. Note, however, that we can, e.g., treat a Luenberger observer by introducing it as a state component $\hat{\mathbf{y}}$ adjoined to each $\mathcal{X}_j, \mathcal{Y}_j$ with suitably defined dynamics involving the current sensor values and $id_{\hat{\mathbf{y}}}$ adjoined to each Y_j .

One reason to exclude Zeno phenomena is to avoid potential difficulties with a recursive use of (S)-d,e. in constructing solutions. In the proof of Theorem 2 below, the positive switching costs $c(j \curvearrowright j')$ enforce this exclusion automatically in optimization. Such costs are a well-known practical reality (often corresponding to the residual effects of rapid scale transients in chattering; compare [8, 2]). It is therefore a common practice to introduce *dead time* following (some) transitions $k \curvearrowright k'$, temporarily preventing a repetition of the mode k or of that transition. We can include such dead time in the formulation by introducing a state component $z \in \mathcal{Z} = \mathbb{R}^J$, satisfying $z_{k \curvearrowright k'} = 1$ for the relevant switching indices $(k \curvearrowright k') \in \mathcal{J}' \subset \mathcal{J} \times \mathcal{J}$, and adjoining \mathcal{Z} to each $\mathcal{X}_j, \mathcal{Y}_j$; as part of $\mathbf{f}(\cdot; k \curvearrowright k')$ one then resets $z_{k \curvearrowright k'}$ to 0. Now one obtains the dead time effect by deleting k from each $\mathcal{A}_j(y, z)$ or deleting k' from $\mathcal{A}_k(y, z)$ until $z_{k \curvearrowright k'}$ reaches its threshold. Note that being able to treat this kind of resetting of z to handle dead time is now our principal reason for retaining a transition map \mathbf{f} as part of (E).

There are also problems for which we do expect the possible occurrence of such behavior. It may then be convenient to view this behavior as a whole, defining within our framework a *chattering mode* (or idealized as a *sliding mode*). \square

Without further hypotheses it is easy to construct examples for which no global solutions exist at all: for example, the switching rules as given might produce a sequence of switching times $t_\nu \nearrow t_* < T$ — violating **(S)**-b. and also with no way to obtain a continuation after t_* . Also, since $\mathcal{A}_j(y)$ need not be a singleton, we cannot expect that **(S)** will determine solution evolution uniquely when solutions do exist. Typically the sets $\mathcal{C}^{j \curvearrowright j'}$ will be *switching surfaces* with the trajectories transverse to these and \mathbf{y} leaving \mathcal{S}_j so switching is forced. We must, however, allow for the alternative possibility that \mathbf{y} , continuing from $y \in \mathcal{C}^{j \curvearrowright j'}$ using mode j , would remain (at least briefly) in \mathcal{S}_j and the choice would be genuine. Such *anomalous points* are a major technical difficulty for this theory and we will further discuss their effect later, noting here only that this is an inherent source of non-uniqueness for solutions since we will be accepting *both* possible choices as legitimate. In the next sections we re-examine the formulation above in the light of possible implementation and impose hypotheses ensuring existence.

3. Modeling and interpretation

Mathematical models are always created, selected, and analyzed with a purpose and we keep this functionality at the forefront of our present concern: convenience is one of the major desiderata in the selection of appropriate models. While control theory is inherently a prescriptive approach to the world in which we may be inclined to ignore the descriptive aphorism, “Natura non facit saltus” (“*Nature does not make jumps*,” attributed to Newton, Leibniz, Linnaeus, . . .), we recognize that any control design is useful only as implemented:

A prescriptive model should be a descriptive model of its implementation

so we must have some concern that the nominal behavior of these discontinuous systems is consistent with their actual behavior. In this section we complement the prescriptive formulation **(E)**, **(S)** with some interpretive comments on the construction from this point of view, clarifying our choices of assumptions.

The fundamental principle of such interpretation is that hybrid systems are a simplified description of multiscale problems in which the transitions $j \curvearrowright j'$ which we are describing as ‘instantaneous’ are actually taking place on a faster time scale than we wish to model; see, e.g., [10]. [If $\mathcal{X}_{j'} = \mathcal{X}_j$, the transition function $\mathbf{f} : \mathbf{x}(t-) \curvearrowright \mathbf{x}(t+)$ might then simply reflect the result of state evolution on the rapid scale.] As with any modeling, success means that we have taken into account those aspects whose effects are inescapable without treating details which can be ignored. Since our description is then an idealization of the world, we are inclusive in the consideration of mild solutions, so our version of well-posedness will require that

The limit of solutions will itself be accepted as a solution,

i.e., the solution set depends upper semicontinuously on the data.

Note that we are permitting degenerate interswitching intervals \mathcal{I}_ν with $t_{\nu-1} = t_\nu$. This might simply correspond to the possibility, which we want to

include here, that distinct effects can occur simultaneously on the modeling scale, meaning only that we cannot determine priority without resolving aspects of the rapid behavior which we are content to leave hidden from us; we do insist that the sequencing, particularly that of the associated modes j_ν , be preserved since this priority may be significant in determining the subsequent evolution on our modeling scale. In such a situation we cannot predict the outcome definitively with the information available. On the other hand, by accepting the alternatives as equally valid solutions we are able to say that,

“What happens must be one of these possibilities.”

(to within the level of approximation corresponding to the usual model uncertainty). An arbitrary selection might provide uniqueness, but lacking a selection principle justifiable from considerations of the unknown rapid behavior we are primarily concerned not to exclude any genuine possibility and so reject such an artificial uniqueness as spurious. This is done in much the same spirit as the acceptance of ‘weak’ or ‘mild’ or ‘generalized’ solutions since at worst these are idealized versions of genuine possibilities and this idealization may not permit us the luxury of restricting our attention to ‘classical solutions.’ We will refer to the times and the situations giving this ambiguity as *anomalous points*. [A related possibility would be a cascade with several transitions $j \curvearrowright j' \cdots \curvearrowright \bar{j}$ occurring as a sequence on the fast scale; it is always possible, but perhaps inconvenient, to replace this by an equivalent compound single switching event $j \curvearrowright \bar{j}$.]

In view of the above, $\mathbf{j}(\cdot)$ need not be a ‘function’ on $[0, T]$ in the usual sense. However, we can think of it simply as a finite *modal sequence* of pairs $(j, \tau)_\nu \in \mathcal{J} \times \mathbb{R}_+$ with τ_ν the length of the ν -th interswitching interval \mathcal{I}_ν so $\sum_\nu \tau_\nu = T$; one recovers the switching times as $t_\nu = \tau_1 + \cdots + \tau_\nu$ and recovers $\mathbf{j}(t)$, when t is not a switching time, by **(S)**-c. Abusing notation somewhat, we continue to denote these by $\mathbf{j}(\cdot)$. We topologize the set $\mathcal{MS}[0, T]$ of all such modal sequences on $[0, T]$ as follows:

Definition 3.1. $[\mathbf{j}^m \rightarrow \mathbf{j}]$ in $\mathcal{MS}[0, T]$ means that each $j_\nu^m \equiv j_\nu$ for large m and each $\tau_\nu^m \rightarrow \tau_\nu$ in \mathbb{R}_+ subject to the constraint $\sum_\nu \tau_\nu^m = T$.

Somewhat similarly, a solution $\mathbf{x}(\cdot)$ would not be a ‘function’ on $[0, T]$ even if there were no change in state spaces: we retain, at any switching time t , both values $\mathbf{x}(t-), \mathbf{x}(t+)$ and, even in contexts with degenerate interswitching intervals, include both when discussing a corresponding *trajectory* $[[\mathbf{x}]] = \{\mathbf{x}(t) : t \in [0, T]\}$. We view the switching as occupying time on a more rapid scale — so the transition map f might represent evolution on that rapid scale — but we make no attempt to include more of the course of this evolution as a connecting part of the trajectory. With this treatment we note, from the continuity of each π_j , that the trajectory $[[\mathbf{x}]]$ for any solution $\mathbf{x}(\cdot)$ on any $[0, T]$ will be compact in $\cup_j \mathcal{X}_j$.

On the other hand, there might be a still slower time scale on which the switchings we are here describing become a rapidly repetitive *chattering mode*, averaging as a *sliding mode*, switching infinitely often within a finite period. These

situations are certainly important and have been treated extensively (cf., e.g., [3, 11, 1, 8]), but they are not our present concern and, as is essential for our treatment here, we will adopt hypotheses bounding the number of pairs in any modal sequence as above for any bounded period $[0, T]$; compare **(S)**-b. which forbids Zeno phenomena for feedback solutions.

We turn now to considering the *open loop problem* in which a fixed modal sequence \mathbf{j} is specified as data. [We continue to use **(E)**, **(S)**, but note that **(E)**-b., c. are here irrelevant: effectively we are taking each \mathcal{A}_j independent of y in defining ‘admissibility’ of \mathbf{j} , so Y_j is not needed.]

Theorem 1. *Let an admissible $\mathbf{j}(\cdot)$ be given as data and suppose suitable initial data ξ given in \mathcal{X}_{j_1} . Then there is a unique solution of the open loop problem specified by **(E)**, **(S)** and this depends continuously, in an appropriate sense, on the specified \mathbf{j} and ξ .*

PROOF: Existence is immediate, recursively constructed uniquely by alternately using **(S)**-d., e. starting with $\mathbf{x}(0) = \xi$ and $\mathbf{x}(t) = \pi_{j_1}(t, 0, \xi)$ on $\mathcal{I}_1 = [0, t_2]$, etc., so we need only verify continuous dependence. Our definition of convergence $\mathbf{j}^m \rightarrow \mathbf{j}$ means that only the interswitching times τ_ν^m change with m so, recalling the assumed continuity of the transition maps and dynamical systems involved, the same recursion also shows that $\mathbf{x}^m(t_\nu^m \pm) \rightarrow \mathbf{x}(t_\nu \pm)$ (even taking $\xi^m \rightarrow \xi$ and even if some interswitching intervals become degenerate in the limit). We similarly get $\mathbf{x}^m(t) \rightarrow \mathbf{x}(t)$ for any t in the interior of an interswitching interval for \mathbf{j} and assume that any ‘appropriate sense’ for convergence of the solutions will follow from this, e.g., we exclude the use of an L^∞ topology for solutions. \square

Remark 3.2. The statement and proof above are ambiguous as to the total interval but we may think of this as finite $[0, T]$ and, as usual with T arbitrary, this also provides the result on $[0, \infty)$.

We now set

$$\begin{aligned} \mathcal{MS}^N &= \{\mathbf{j} \in \mathcal{MS}[0, T] : \text{there are at most } N \text{ switches}\}, \\ \mathcal{K}^N(\xi) &= \{[\mathbf{x}] : \mathbf{x}(\cdot) \text{ corresponds to } \mathbf{j} \in \mathcal{MS}^N, \mathbf{x}(0) = \xi\}. \end{aligned}$$

It is easy to see that each of the subsets $\mathcal{MS}^N[0, T]$ will be compact in $\mathcal{MS}[0, T]$. We have already noted that each individual trajectory $[\mathbf{x}]$ is compact and, from the discussion of continuous dependence in the proof above, we now see that each $\mathcal{K}^N(\xi)$ is compact. \square

Still in the setting of the open-loop problem, but now in a context of infinite horizon optimal control, we consider choice of the modal sequence so as to minimize a cost functional of the form

$$\Psi = \int_0^\infty e^{-\beta t} c_{\mathbf{j}(t)}(\mathbf{x}(t)) dt + \sum_{\nu=2}^\infty e^{-\beta t_\nu} c(j_{\nu-1} \curvearrowright j_\nu). \quad (3.1)$$

We wish to show that the inf defining the *value function*

$$V_j(\xi) = \inf\{\Psi[\mathbf{j}, \xi] : \mathbf{j}(0) = j, \mathbf{x}(0) = \xi\} \quad (3.2)$$

is actually an attained minimum.

Theorem 2. *Assume that each running cost $c_j(\cdot) \geq 0$ of (3.1) is continuous; suppose $j_1 = j$ and suitable initial data ξ are given in $\mathcal{J}, \mathcal{X}_j$. Let each switching cost $c(j \curvearrowright j') > 0$ and assume there is some \mathbf{j}_* for which Ψ is finite. Then there is a modal sequence (switching control) $\mathbf{j} = \mathbf{j}_*$ for which $\Psi = \Psi[\mathbf{j}, \xi]$ attains its minimum $V_j(\xi)$. This minimum cost depends lower semicontinuously on $\xi \in \mathcal{X}_j$.*

PROOF: The set $\{\mathbf{j} : \Psi < \infty\}$ is nonempty by assumption so we can consider a minimizing sequence \mathbf{j}^m : $\Psi^m = \Psi[\mathbf{j}^m, \xi] \rightarrow \inf\{\Psi\} = V_j(\xi)$. For arbitrary $T < \infty$, the switching costs then ensure a bound on the number of transitions during $[0, T]$ so we may extract a convergent subsequence; further extracting subsequences we can assume $\mathbf{j}^m \rightarrow \mathbf{j}^*$ on every bounded interval. Theorem 1 applies to the problem on each $[0, T]$, showing the corresponding solutions converge $\mathbf{x}^m \rightarrow \mathbf{x}^*$ there ‘in a suitable sense.’ From the form of (3.1) we easily see this implies convergence of the restricted costs:

$$\Psi^m \Big|_{[0, T]} \rightarrow \Psi^* \Big|_{[0, T]} \quad \text{so} \quad \Psi^* \Big|_{[0, T]} \leq \Psi^m \Big|_{[0, T]} + \varepsilon \leq \Psi^m + \varepsilon \rightarrow V_j(\xi).$$

Letting $T \rightarrow \infty$, this shows that $\Psi^* \leq V_j(\xi)$ so $V_j(\xi)$ is a min with minimizer \mathbf{j}^* . If \mathbf{j}^m is the minimizer for $\xi^m \rightarrow \xi$, then we can extract a convergent subsequence as above to get $\mathbf{j}^m \rightarrow \mathbf{j}^*$ and see

$$V_j(\xi) \leq \Psi[\mathbf{j}^*, \xi] \leq \liminf_m \Psi[\mathbf{j}^m, \xi^m] = \liminf_m V_j(\xi^m). \quad \square$$

[It is not difficult to see that $V_j(\cdot)$ is actually continuous if each π_j is locally uniformly continuous.]

4. Modeling feedback

Suppose we consider the optimization problem of Theorem 2 for autonomous dynamical systems so autonomy of the system makes the value function V independent of any starting time and

$$V_j(\xi) = \Psi^* \Big|_{[0, \tau]} + e^{-\beta\tau} V_{\mathbf{j}^*(\tau)}(\mathbf{x}^*(\tau)) \quad (4.1)$$

for each $\tau > 0$, where $\mathbf{j}^*, \mathbf{x}^*$ are optimal as in the proof of Theorem 2. We would like to recover the optimal switching control from V , allowing for the possibility that this need not be unique. The possibility of a transition $j \curvearrowright j' \neq j$ when $\mathbf{x}(t) = \xi$ just means that j' is in \mathcal{A}_j and

$$\left[\begin{array}{l} \text{some optimal } \mathbf{j} \text{ starting at } (j, \xi) \\ \text{immediately switches } j \curvearrowright j' \neq j \end{array} \right] \quad \text{---} \quad (4.2)$$

$$c(j \curvearrowright j') + V_{j'}(\mathbf{f}(\xi; j \curvearrowright j')) = V_j(\xi)$$

where equality just means that the optimal value can be attained with a switch to j' . On the other hand, comparing with (4.1) and (3.1), we see that

$$\left[\begin{array}{c} \text{some optimal } \mathbf{j} \text{ starting at } (j, \xi) \\ \text{continues in mode } j \end{array} \right] \quad \text{---} \quad (4.3)$$

$$\int_0^\tau e^{\beta t} c_j(\pi_j(t, \xi)) dt + e^{\beta \tau} V_j(\pi_j(\tau, \xi)) = V_j(\xi) \quad (\text{some } \tau > 0).$$

Remark 4.1. From this we observe that:

Let $\mathcal{J}, \mathbf{f}, \pi_j$ be as in Theorem 2; assume each π_j is autonomous. Set $\mathcal{Y}_j = \mathcal{X}_j$, $Y_j = \text{id}(\mathcal{X}_j)$, and

$$\mathcal{A}_j(\xi) = \{j \text{ if (4.3), } j' \text{ if (4.2)}\} \quad (4.4)$$

for $j \in \mathcal{J}$, $\xi \in \mathcal{X}_j$ to complete the specification **(E)**. Let (\mathbf{j}, \mathbf{x}) , starting with (j_1, ξ) , be as in Theorem 1.

Then the pair (\mathbf{j}, \mathbf{x}) is optimal for the switching control problem of Theorem 2 if and only if it is a feedback solution as in **(S)**.

PROOF: Clearly any optimal control satisfies **(S)** with (4.4). Conversely, by connectedness and the continuity of $\mathbf{y}(\cdot) = \mathbf{x}(\cdot)$, such a solution of **(S)** satisfies (4.1) on each nondegenerate interswitching interval \mathcal{I}_ν (hence) and on $[0, T]$ by induction on ν , hence is optimal. \square

This is a primary motivation for taking **(S)** as defining the general structure of feedback we consider here, while noting, for example, that we cannot always expect to have full-state feedback as in Remark 4.1 and would necessarily implement only finitely many sensors. Thus, we consider the evolution of a solution for **(S)** as an independent problem, with the elements of **(E)** somewhat general. Purely for expository convenience, however, we assume henceforth that $\mathcal{X}_j, \mathcal{Y}_j, Y_j$ are each independent of $j \in \mathcal{J}$ and that the dynamical systems π_j are autonomous.

Recall that the sensor maps Y_j and the resulting sensor output $\mathbf{y}(\cdot)$ played no role in Theorems 1 and 2, but the regularity to be expected of these is now a significant concern in being able to evaluate \mathbf{y} pointwise in t so the conditions of **(S)** make sense. This regularity and its interaction with the avoidance of Zeno behavior — i.e., with **(S)**-b. — constitute the essential technical difficulties in analyzing this feedback structure. For Theorem 1, **(S)**-b. was already an admissibility hypothesis on the given \mathbf{j} and in the proof of Theorem 2 this was a consequence of the assumed positivity of the switching costs. For a general feedback we will need new hypotheses; we begin by assuming the feedback diagram and sensor map satisfy the following set of hypotheses.

- (H₁)
- a. each $\mathcal{C}^{j \curvearrowright j'}$ is closed in \mathcal{X} and $\mathcal{S}^j \supset [\mathcal{Y} \setminus \mathcal{B}^j]$.
 - b. Y is set-valued with $Y(\xi)$ finite and nonempty for each $\xi \in \mathcal{X}$.
 - c. Y is upper-semicontinuous, i.e., if one has $y_k \in Y(x_k)$ with $x_k \rightarrow \bar{x}$ in \mathcal{X} , then there is a subsequence $(y_{k(\ell)})$ converging to some $\bar{y} \in Y(\bar{x})$.
 - d. cascades of the form $j \curvearrowright j$ are forbidden: i.e., there exists no sequence of pairs $(j, \xi)_{\nu=1}^{\bar{\nu}}$ with $j_1 = j_{\bar{\nu}}$ such that

$$Y(\xi_{\nu}) \cap \mathcal{C}^{j_{\nu} \curvearrowright j_{\nu+1}} \neq \emptyset, \quad \xi_{\nu+1} = \mathbf{f}(\xi_{\nu}; j_{\nu} \curvearrowright j_{\nu+1}).$$

It is precisely at this point that our considerations will depend in an essential way on the particular PDE setting since we have in mind, at least as an idealization, that our sensors will be point evaluations in the spatial domain of (1.1). For the operation of a thermostat, where (1.1) becomes a heat equation, one has more than enough regularity that this causes no difficulty (provided the sensor location is separated from the furnace/AC). For a transport equation, however, the occurrence of modal switching can be expected to introduce spatial discontinuities which propagate to the sensors and cause temporal discontinuities in $\mathbf{y}(\cdot)$; it then becomes a delicate problem (cf. [5]) to provide a space \mathcal{X} which allows for this and at the same time gives both continuity of the dynamics and adequate regularity of $\mathbf{y}(\cdot)$.

We now provide an additional hypothesis which, along with (H₁), will suffice to give (S)-b. in showing the existence of solutions for the feedback problem. This hypothesis (H₂) is rather technical, but, as an example, we will later show how to verify these hypotheses for transport on a graph.

- (H₂)
- There exists $\bar{\tau} > 0$ such that for each $\xi \in \mathcal{X}$, $T \geq \bar{\tau}$, and N' there exists $N = N(\xi, N', T)$ such that:
- if $T - \bar{\tau} < T' < T$ and $\mathbf{j} \Big|_{[0, T - \bar{\tau}]}$ is in $\mathcal{MS}^{N'}$, then there are no more than N points of $\bar{\Xi} = \overline{\{\xi \in \mathcal{X} : \#Y(\xi) \neq 1\}}$ in the trajectory $\{\mathbf{x}(t) : t \in [0, T']\}$.

[While we have formulated this hypothesis to obtain a context of piecewise continuous $\mathbf{y}(\cdot)$, one might expect that a rather similar treatment could be formulated for, e.g., $\mathbf{y}(\cdot)$ of bounded variation.]

Theorem 3. *Assume we have (E) satisfying (H₁), and (H₂). Then, for any given $(j, \xi) \in \mathcal{J} \times \mathcal{X}$, there is $[\mathbf{j}(\cdot), \mathbf{x}(\cdot), \mathbf{y}(\cdot)]$, a global solution of the feedback problem starting with (j, ξ) .*

PROOF: It is convenient to restrict our attention to ‘skittish solutions,’ which switch whenever that is allowable under the switching rules of (S). By Zorn’s Lemma one has existence of a maximally defined skittish solution $[\mathbf{j}(\cdot), \mathbf{x}(\cdot), \mathbf{y}(\cdot)]$

whose domain necessarily has one of the forms $[0, 0]$, $[0, T_*]$, $[0, T_*)$, or $\mathbb{R}_+ = [0, \infty)$; we wish to show this can only be $[0, \infty)$.

From (2.3), we have initially either $\mathbf{y}(0) \cap \mathcal{B}^j \neq \emptyset$ and proceed with a maximal finite cascade $j \curvearrowright \bar{j}$ or have $\mathbf{y}(0) = Y(\xi)$ in $\mathcal{S}^j \setminus \mathcal{B}^j$. Since (\mathbf{H}_1) -a. gives $\mathcal{S}^j \setminus \mathcal{B}^j = \mathcal{Y} \setminus \mathcal{B}^j$ open and (\mathbf{H}_1) -c. ensures a solution can remain for some (small) interswitching interval in mode j . In the former case, the cascade ends with $j' \curvearrowright \bar{j}$ leaving $\mathbf{x} = \bar{\xi}$ with $\mathbf{y} = \bar{\eta} = Y(\bar{\xi}) \notin \mathcal{B}^j$ (or the cascade could have continued); by (2.3) we then have $\bar{\eta} \in \mathcal{S}^j$ and the solution could be extended. In either case, then, the domain $[0, 0]$ is inconsistent with maximality. Similarly, a domain $[0, T_*]$ is also inconsistent with maximality since we could restart the problem at T_* and use the same argument.

Next suppose the maximal domain were of the form $[0, T_*)$. Since $[\mathbf{j}(\cdot), \mathbf{x}(\cdot), \mathbf{y}(\cdot)]$ is a solution on every subinterval, either there is a last switching $\cdot \curvearrowright j_*$ at $t_* < T_*$ or the sequence of switching times (t_ν) converges to T_* , violating (\mathbf{S}) -b. on $[0, T_*]$ itself. In the former case, $t \mapsto \mathbf{x}(t) = \pi_{j_*}(t - t_*, \mathbf{x}(t_*+))$ is continuous on $[0, T_*]$ and either $\mathbf{y}(T_*) = Y(\mathbf{x}(T_*)) \in \mathcal{S}^{j_*}$ — so the solution continues through T_* in mode j_* by (\mathbf{H}_1) -a. — or $\mathbf{y}(T_*) \cap \mathcal{B}^{j_*} \neq \emptyset$ so one can switch and the solution can be extended at least to $[0, T_*]$; either of these possibilities contradicts the maximality of $[0, T_*)$.

In the latter case, with $t_\nu \rightarrow T_*$, the maximally defined $\mathbf{j}(\cdot)$ necessarily consists of an infinite sequence of nondegenerate interswitching intervals of length $\tau_\mu > 0$ (with $\sum_{\mu=1}^\infty \tau_\mu = T_*$) separated by maximal cascades $j_{\mu-1} \curvearrowright j_\mu$. Choose any $T' \in (T_* - \bar{\tau}, T_*)$, let N' bound the number of switchings in $\mathbf{j}(\cdot)$ on $[0, T']$, and set $N = N(\xi, N', T_*)$ as in (\mathbf{H}_2) . Now consider any one of the interswitching intervals $\mathcal{I} = \mathcal{I}_\mu = [t', t'']$ (i.e., $t' = t_{\mu-1}$, $t'' = t_\mu$) with $t' \geq T'$ on which $\mathbf{j} \equiv j = j_\mu$. By (\mathbf{S}) -c, this must be initiated with $\mathbf{x}(t'-) = \xi^1$ producing a maximal cascade $j_{\mu-1} = j^1 \curvearrowright \cdots \curvearrowright j^n = j$ with $Y(\xi^\nu) \in \mathcal{C}^{j^\nu \curvearrowright j^{\nu+1}}$ and $\xi^{\nu+1} = \mathbf{f}(\xi^\nu; j^\nu \curvearrowright j^{\nu+1})$ for $\nu = 1, \dots, n-1$ as in (\mathbf{H}_1) -d. Assuming no points of Ξ occur in this sequence (or during \mathcal{I}) so Y is simply a continuous single-valued function there, one can show easily that the set S of points in \mathcal{K} which can initiate this particular sequence (as ξ^1) is closed and in $\mathcal{K}^{N'}$, so compact in \mathcal{X} . Thus, iterating \mathbf{f} , the set S' of points terminating the sequence (as $\xi^n \xi_+$) is also compact and $S'' = Y(S')$ is compact in \mathcal{Y} — with $S'' \cap \mathcal{B}^j = \emptyset$, as the cascade is maximal. Hence there is a minimal distance from S'' to \mathcal{B}^j . We must have $\mathbf{y}(t'') \in \mathcal{B}^j$ to end \mathcal{I} by initiating another transition and note that $[t \rightarrow Y(\pi_j(t - t', \xi_{n+}))]$ is uniformly continuous on \mathcal{I} so there is a minimal time required to make this transit; with only finitely many possibilities for the cascade, this time τ_* may be taken as the same for all so the length τ_ν of such an interswitching interval is bounded below by τ_* and there can be at most $\bar{\tau}/\tau_*$ such intervals. We have no lower bound on the length of those interswitching intervals involving points of Ξ , but the number of these is bounded by our technical hypothesis (\mathbf{H}_2) , contradicting the assumption above of an infinite sequence $\{\mathcal{I}_\mu\}$.

Thus, the maximal domain must be $[0, \infty)$; as desired, the maximally defined skittish solution is global. Of course, this need not be unique and there may also be additional (non-skittish) global solutions. Note also, from this proof, that if we are given, on a bounded domain, any $[\mathbf{j}(\cdot), \mathbf{x}(\cdot), \mathbf{y}(\cdot)]$ satisfying **(S)** there, then it can be extended to a global solution. \square

Example 4.2. As a first example, consider a thermostat-controlled heating system. For the simplest case, one would have a single point-evaluation sensor: $Y : \xi \mapsto \eta = \xi(p)$ with p given in the spatial region Ω and $\xi \in \mathcal{X} = C(\Omega)$. We take the effect of the control in the boundary flux so (1.1) becomes the heat equation for the temperature distribution $\mathbf{x}(t, \cdot)$

$$\mathbf{x}_t = \Delta \mathbf{x} \text{ on } \Omega, \quad \mathbf{x}_\nu = \alpha \mathbf{x} + v_j \text{ at } \partial\Omega \quad (4.5)$$

defining π_j for the two modes $j \in \mathcal{J} = \{0, 1\}$ denoting OFF/ON. [Here the flux difference $v_1 - v_0$ gives the effect of the furnace or AC.] We have no jumps in the state itself when the thermostat switches so $\mathbf{f} = id_{\mathcal{X}}$. The well-posedness of (4.5) is standard and **(H₁)**-c.,d. as well as **(H₂)** are immediate since Y is single-valued and continuous.

Now let η_* be our setpoint, the desired temperature, and allow a margin $\pm\delta$ with $\delta > 0$. Then switching is determined by

$$\begin{aligned} \mathcal{A}_0(y) &= \begin{cases} \{0\} & \text{if } y > \eta_* - \delta \\ \{0, 1\} & \text{if } y = \eta_* - \delta \\ \{1\} & \text{if } y < \eta_* - \delta \end{cases} & \mathcal{A}_1(y) &= \begin{cases} \{0\} & \text{if } y > \eta_* + \delta \\ \{0, 1\} & \text{if } y = \eta_* + \delta \\ \{1\} & \text{if } y < \eta_* + \delta \end{cases} \\ \text{so} \quad \mathcal{C}^{0 \wedge 1} &= (-\infty, \eta_* - \delta], \quad \mathcal{S}^0 = [\eta_* - \delta, \infty), \\ \mathcal{C}^{1 \wedge 0} &= [\eta_* + \delta, \infty), \quad \mathcal{S}^1 = (-\infty, \eta_* + \delta]. \end{aligned}$$

I.e., the furnace turns ON when temperature (at the thermostat) falls below $\eta_* - \delta$ and goes OFF when it rises above $\eta_* + \delta$. [The resulting transducer: $y(\cdot) \mapsto \mathbf{j}(\cdot)$ is precisely the hysteretic *non-ideal relay* of [6, section 28.2], well defined except for the possible ambiguity of anomalous points.] We have **(H₁)**-a.,c.,d. trivially; with $\delta > 0$, **(H₁)**-b. holds as $\mathcal{C}^{0 \wedge 1} \cap \mathcal{C}^{1 \wedge 0} = \emptyset$, and **(H₁)**-e. holds as $y(\cdot)$ is continuous here with each $\mathcal{S}^j \setminus \mathcal{B}^j$ open.

Taking $\delta > 0$ is implicit in the usual design of thermostats and we note that our hypotheses fail for the idealized thermostat with $\delta = 0$. In that setting one has a (pointwise) functional map: $y \mapsto j$ and convexifying when $y = \eta_*$ (compare [4, 3]) one does obtain existence, although with the possibility of Zeno-ness in the form of sliding modes: ON/OFF oscillation of the furnace on the rapid scale. \square

Example 4.3. We conclude with a more demanding example, considering transport on a graph with feedback modal control: descriptively, we imagine reacting chemical species being transported by a solvent, moving as plug flow along the pipe segments $\{E_m : m \in \mathcal{M}\}$ of a network. These single-segment problems are then coupled at each node N_n of the resulting graph Γ through the allocation of incoming flux, including exogenous sources, to outgoing segments, including external outputs). Our presentation here largely follows the more detailed treatment

in [5].

The state $\mathbf{x}(t)$ in this example will be the densities (concentrations) $u(t, \cdot)$ of conserved species of interest, taken in a suitable state space \mathcal{X} of vector functions on Γ . The map $Y : \mathcal{X} \rightarrow \mathcal{Y} = \mathbb{R}^K$ is given by evaluations $y_k = u_{i(k)}(\cdot, \bar{s}_k)$ at specified sensor points $\bar{s}_k \in \Gamma$. We have $\mathbf{f} = id_{\mathcal{X}}$ here and initially let the feedback diagram be subject only to (2.3), (\mathbf{H}_1) -a., and $\mathcal{C}^{j \wedge j'} \cap \dots \cap \mathcal{C}^{j'' \wedge j} = \emptyset$, which is here equivalent to (\mathbf{H}_1) -b.

For simplicity of exposition we assume an incompressible carrier (solvent) and uniform cross-sectional area α_m in each pipe segment E_m and input end 0_m independent of the mode; the transport is produced by the action (specified by j) of a pump at 0_m . The flow velocity v_m^j will then be constant on E_m and, again for simplicity we assume v_m^j is also constant in t . The evolution π_j of the system is now determined by these flow velocities. First, we have a set of convection/reaction equations: on each of the individual edges

$$u_t + v_m^j u_s = f(u) \quad \text{on } E_m \quad (4.6)$$

and will use the classical *method of characteristics* to construct solutions:

Let $\omega(t; \omega_0)$ be the solution of the ordinary differential equation

$$\omega' = f(\omega), \quad \omega(0) = \omega_0, \quad (4.7)$$

Given (t, s) , track back along the characteristic $\sigma(\tau) = s - [t - \tau]v$ to an initialization point (τ_0, σ_0) — either $\tau = \tau_0 \leq t$ is a starting time (i.e., 0 or the most recent switching time) with $\sigma_0 = \sigma(\tau_0) \in E_m$ or else $\sigma_0 = 0_m$ with $\tau_0 \leq \tau \leq t$. Now set $u(t, s) = \omega(t - \tau; \omega_*)$ where ω_* is the given data at (τ_0, σ_0) .

The construction of π_j is then completed by the nodal coupling, specifying the input data $u_{*m}(\cdot)$ to each pipe. For each node N_n we have input edges \mathcal{M}_n^+ and output edges \mathcal{M}_n^- (with $\cup_n \mathcal{M}_n^- = \mathcal{M} = \cup_n \mathcal{M}_n^+$). Clearly the assigned flow velocities must satisfy the consistency condition

$$[\text{flux in}]_n^j = \sum_{m \in \mathcal{M}_n^-} \alpha_m v_m^j = \sum_{m \in \mathcal{M}_n^+} \alpha_m v_m^j = [\text{flux out}]_n^j = \Phi_n^j \quad (4.8)$$

Assuming perfect mixing at the node, the vector of combined input concentrations at N_n of the chemical species will be

$$U_n(\tau) = \frac{\sum \{\alpha_m v_m^j u_m(\tau, 1_m) : m \in \mathcal{M}_n^-\}}{\sum \{\alpha_m v_m^j : m \in \mathcal{M}_n^-\}} \quad (4.9)$$

and the required input data to E_m is then given by

$$u(\tau, 0_m) = u_*(\tau) = U_n(\tau) \quad \text{for } m \in \mathcal{M}_n^+ \quad (4.10)$$

[This must be modified in the case of exogenous sources, for which one can permit some choice in the formulation.]

We take this construction along characteristics as defining our notion of solution for (4.6) and so the definition of π_j .

Our major technical difficulty in this example is to specify and topologize the state space \mathcal{X} so as to verify the hypotheses **(H₁)** and **(H₂)** while maintaining continuity of this π_j . From our solution construction and the continuity of $\omega(\cdot, \cdot)$, we see that discontinuities will propagate along the characteristics (including across nodes) and can be created only at nodes at switching times. We expect, then, that the state $\mathbf{x}(t) = u(t, \cdot)$ will be a piecewise continuous function and will take $\mathcal{X} = \mathcal{P}$ to be a suitable space of such functions.

As with modal sequences $\mathbf{j}(\cdot)$, it is possible to create degenerate ‘intervals of continuity’ — allowing $u(t, \cdot)$ to be continuous on an interval $[s, s']$, take a value on the degenerate interval $[s', s'']$ with $s'' = s'$, and then again continuous on $[s'', s''']$. This could occur if discontinuities propagating through edges E_m and $E_{m'}$ incoming to the same node N_n arrive simultaneously; in view of our modeling considerations we interpret ‘simultaneously’ as meaning ‘indistinguishably close’ — although possibly distinct on the rapid time scale so we retain both possibilities with the alternative intermediate values. These degenerate intervals correspond to a fine spatial scale, comparable to the rapid time scale. In view of this possibility we must be careful with the interpretation of the sensor map Y , taking this to be set-valued when such a subinterval coincides with one of the sensor points \bar{s}_k .

This suggests our characterization of an element of $\mathcal{X} = \mathcal{P}$: for each m one has a vector-valued piecewise continuous functions on closed subintervals, including possible finite sequences of degenerate subintervals as with \mathcal{MS} and then, much as with Definition 3.1, we topologize this as follows:

Definition 4.4. $[u^k \rightarrow u]$ in \mathcal{P} if, for each E_m , the number of subintervals is eventually fixed, the dividing endpoints converge, and the functions on them (normalized to domain $[0, 1]$ with values on degenerate subintervals taken as constants) converge in the sense of $C[0, 1]$.

One easily sees that the problem is well-posed in this setting: π_j is continuous from $\mathbb{R}_+ \times \mathcal{P}$ to \mathcal{P} . As suggested earlier, we use point evaluations to define $Y : \mathcal{P} \rightarrow \mathcal{Y} = \mathbb{R}^K$ by

$$Y(\xi) = [\xi_{i(1)}(\bar{s}_1), \dots, \xi_{i(K)}(\bar{s}_K)] \quad (\xi \in \mathcal{P}) \quad (4.11)$$

with the provision that: if a discontinuity of ξ occurs at one of the sensor points \bar{s}_k so $\xi(\cdot)$ has both left- and right-hand values there (perhaps even more values if this involves a degenerate subinterval), then $y_k(\xi)$ becomes the set of all relevant values; this clearly gives **(H₁)**-b.,c. It is not difficult to construct the feedback diagram to give **(H₁)**-a.,d. — e.g., taking \mathcal{S}^j to contain the open set $\mathcal{Y} \setminus \mathcal{B}^j$, perhaps adjoining (as anomalous points) other points $\eta \in \mathcal{B}^j$ for which $Y^{-1}(\eta)$ contains ξ from which one might wish to extend the solution in mode j — and we assume this.

In order to satisfy **(H₂)** we require that the sensor points are separated from the actuators, i.e., from the input nodes where discontinuities might be created. Thus, we will assume there is some $\bar{\tau} > 0$ such that

$$[\bar{s}_k - 0_m] / v_m^j \geq \bar{\tau} \quad \text{for all } j \in \mathcal{J}, \bar{s}_k \in E_m, k = 1, \dots, K. \quad (4.12)$$

With this assumption, no discontinuity created after time $t = T$ could possibly be propagated along characteristics to arrive at any sensor before $t = T + \bar{\tau}$. If we have bounded by N' the number of switchings in $\mathbf{j}(\cdot)$ up to $T_* - \bar{\tau}$, then our dynamics and the graph geometry bound both the number of spatial discontinuities arriving to any sensor point, creating a point of Ξ , up to $T_* - \bar{\tau}$ and the number of discontinuities in $\mathbf{x}(T_* - \bar{\tau})$, viewed now as an ‘initial’ state, and so bounds the number which can arrive to a sensor point, creating a new point of Ξ , by any time $T' < T_*$. This total bound is then $N(\xi, N', T)$ and we have verified (\mathbf{H}_2) . \square

References

- [1] M. di Bernardo, C.J. Budd, A.R. Champneys, and P. Kowalczyk, *Piecewise-smooth Dynamical Systems*, (AMS #163) Springer-Verlag, London (2008).
- [2] I. Capuzzo-Dolcetta and L.C. Evans, *Optimal switching for ordinary differential equations*, SIAM J. Cont. Opt. **22**, 1984, pp. 143–161.
- [3] A.F. Filippov, *Differential Equations with Discontinuous Righthand Sides*, Nauka, Moscow (1985) [*transl.* Kluwer, Dordrecht (1988)].
- [4] K. Glashoff and J. Sprekels, *An application of Glicksberg’s theorem to set-valued integral equations arising in the theory of thermostats*, SIAM J. Math. Anal. Appl. **12**, 1981, pp. 477–486.
- [5] F. Hante, G. Leugering, and T.I. Seidman, *Modeling and analysis of modal switching in networked transport systems*, Appl. Math. and Optimization, to appear.
- [6] M.A. Krasnosel’skiĭ and A.V. Pokrovskiĭ, *Systems with Hysteresis*, Nauka, Moscow (1983) [*transl.* Springer-Verlag, Berlin (1989)].
- [7] T.I. Seidman, *Switching systems: thermostats and periodicity* (report **MRR-83-07**), UMBC, 1983. (http://userpages.umbc.edu/~seidman/ss_83.pdf) [modified, as *Switching systems, I*, Control and Cybernetics **19**, 1990, pp. 63–92.]
- [8] T.I. Seidman, *The residue of model reduction*, in *Hybrid Systems III. Verification and Control*, (LNCS #1066; R. Alur, T.A. Henzinger, E.D. Sontag, eds.), Springer-Verlag, Berlin (1996), pp. 201–207.
- [9] T.I. Seidman, *A convection/reaction/switching system*, Nonlinear Anal. – TMA **67**, pp. 2060–2071, 2007.
- [10] T.I. Seidman, *Aspects of modeling with discontinuities*, Int’l. J. Evolution Eqns. **3**, 2008, pp. 129–143.
- [11] V.I. Utkin, *Sliding Modes and their Application in Variable Structure Systems*, Mir, Moscow, 1978.

Acknowledgment

Many thanks to G. Leugering for his support in this work.

Thomas I. Seidman
 Department of Mathematics and Statistics
 University of Maryland Baltimore County
 Baltimore, MD 21250, USA
 e-mail: seidman@math.umbc.edu