Modeling modal transitions in diffusing populations

Thomas I. Seidman*
Department of Mathematics and Statistics,
University of Maryland, Baltimore County,
Baltimore, MD 21250, USA

To Gaston N'Guérékata on the occasion of his 60^{th} birthday

Abstract

We consider a diffusing population as consisting of particles undergoing Brownian motion, each with its own history from a Lagrangian viewpoint. For each such particle we then consider possible "state transitions" determined by crossing thresholds (hysteretic relay as in hybrid systems). Our objective here is to construct a continuum model of the resulting process as a reaction/diffusion system and then to show existence of "solutions" of this system. Technical difficulties arise here in resolving the concerns of hybrid systems (anomalous points and the possibility of Zeno phenomena) in a setting where one is tracking the collective effects on individual diffusing particles without being able to track their individual trajectories. [For visualization, we think of an example of diffusing bacteria and nutrient in which, while undergoing diffusive motion, each bacterium is reacting to its own experience of local nutrient concentration in switching between dormant and active modes.]

*E-mail address: `seidman@umbc.edu`

1 Introduction

Imagine a population of dormant bacteria diffusing in a region Ω.. At some time a supply of nutrient is added locally to part of the region and, where this nutrient concentration $v(x,t_0)$ is above some threshold level η^+, the bacteria are immediately re-activated. [For present purposes we do not question whether instantaneity of bacterial transition between DORMANT and ACTIVE modes is a realistic idealization on a normal time scale, but take it as given that this holds for the bacteria being considered.] At some later time t the nutrient concentration has become $v(\cdot,t)$ and the region may be partitioned into a subregion $\Omega_+(t)$ where $v(x,t) > \eta^+$, another subregion $\Omega_-(t)$ where v is below the threshold η^- for transition to dormancy and an intermediate region $\Omega_*(t) = \{x \in \Omega : \eta^- \le v(x,t) \le \eta^+\}$. What might we expect in $\Omega_*(t)$? If we could have used a Lagrangian viewpoint to follow an individual bacterium with path $z(\cdot)$ such that $z(t) \in \Omega_*(t)$, we would have noted that the bacterium, initially dormant, would have been re-activated at any moment that $z(t)$ crossed the boundary to enter into $\Omega_+(t)$ — by its motion or by a change in Ω_+ — and would then, similarly, have again become dormant at any subsequent moment $z(t)$ crossed the boundary to enter into $\Omega_-(t)$, etc. Thus, the bacterium switches mode ($\pm$ = active or dormant) as it experiences the nutrient concentration $t \mapsto v(z(t),t)$ and its current mode is determined as given by its own most recent transition. Thus, we would have well-determined modes for $\Omega_\pm$, but a bacterium found in Ω_* might be either dormant or active. Modeling this is further complicated by any behavioral differences in these bacterial subpopulations: e.g., active bacteria (reproduce and) metabolize nutrient, but dormant ones do not, so this affects the coupling with evolution of v; they may also diffuse differently.

This example (closely related to the Hoppensteadt-Jäger model described in Section XI.8 of [9]) indicates the kind of situation with which we are primarily concerned. Indeed, the original motivation for the present analysis came from a context suggested by bioremediation and optimal control (compare, e.g., [4] and [7] in which the bacteria remained stationary). While one might suggest a variety of similar problems, the emphasis here will be exclusively on purely mathematical concerns rather than practical consideration of any particular potential application. In particular, we will concentrate on interpreting the formulation so that we can give a positive answer to the fundamental question of the existence of 'solutions'. We will return to these problems in Section 5, but will begin our analysis with the Lagrangian viewpoint, isolating the experience of the individuals in the diffusing population.

When we follow dynamics $z' = \varphi_j(z)$ (for a fixed $j \in M$) we think of this as a "mode". If the *modal index function* $j : [0,T] \to \mathcal{M}$ is piecewise constant, we refer to the jumps from one mode to another as the 'modal transitions' of the title. Thus we would have

$$z' = \varphi_j(z) \qquad\qquad j = j(t) \text{ on } [0,T]. \tag{1}$$

[We do not require autonomous dynamics, but suppress the arguments t in $\varphi_{j(t)}(z(t),t)$.] We will be interested in situations where this switching of the modal index is not specified as input, but is to be determined through the operation of a transducer $\mathbf{W}$ so

$$j(\cdot) = \mathbf{W}[y(\cdot)] \qquad\qquad \text{with } y = Y(v) \tag{2}$$

where v is some auxiliary function whose evolution is coupled with that of z and Y is a suitable scalar function. [For the bacterial example above, we have $y(t) = v(z(t),t)$.] Our concern is with $\mathbf{W}$ whose operation is to switch $j(\cdot)$ from one mode to another when $y(\cdot)$ crosses some threshold, noting that this may be somewhat history dependent so $j(t)$ need not simply be determined by knowing the current value $y(t)$. This kind of modal switching has been extensively studied in connection with hysteresis and in the control-theoretic literature of 'hybrid' or 'variable structure' systems; see also [3], [5], [9], etc. We will be considering descriptively the *nonideal relay* and variations, collecting in Section 2 a review of some relevant material about $\mathbf{W}$. For simplicity of exposition we will mostly restrict our attention to $\mathcal{M} = \{0,1\}$ while noting in Section 2 that it is not difficult to generalize this to considering $\mathcal{M}$ finite.

In considering diffusive contexts, we think both of the effect of stochasticity on a variable structure system and of the coupling of such relay nonlinearities with diffusion equations. In Section 3 we will discuss the formulation of a continuum model of a simple hybrid system with stochastic dynamics and will then introduce a number of complicating variations.

In working with partial differential equations so $z = z(x,t)$, the case where the modal selection is global — think of a thermostat coupled with the heat equation — is easily understood as a hybrid system in which the continuous component has infinite dimensional state space. For our bacterial example, this is not possible and it is then less clear how to formulate the resulting evolution. In Sections **??** and **??** we discuss how this might be related to the stochastic problems when working in the diffusive context and show existence of solutions for a model problem corresponding to that example.

2 Threshold induced switching

In this section we ask how the $\{0,1\}$-valued function $j(\cdot)$ could be determined by a relay from a function $y(\cdot)$, viewed as input. At this point we do not ask how y might be obtained in connection with (2), but take it as given.

For scalar input the 'ideal relay' with threshold η is simply the discontinuous function

$$W_\eta(s) = \begin{cases} 0 & \text{for } s < \eta, \\ 1 & \text{for } s > \eta. \end{cases}$$

Dynamically, this gives the *switching rules* for $\mathbf{W}_\eta : y(\cdot) \mapsto j(\cdot)$ as a transducer:

$$\begin{array}{l} j(\cdot) \text{ switches } 0 \curvearrowright 1 \text{ when } y(\cdot) \text{ increases across } \eta \\ j(\cdot) \text{ switches } 1 \curvearrowright 0 \text{ when } y(\cdot) \text{ decreases across } \eta. \end{array} \tag{3}$$

Note that this mandates switching when $y(\cdot)$ crosses η and we take it also as forbidding switching other than at $y = \eta$. Effectively, this amounts to working with a single "mode" which would be discontinuous across η; cf., e.g., [2].

More interesting (and more typical in control-theoretic applications) is the 'nonideal relay' $\mathbf{W}$ with separated switching thresholds ($\eta^- < \eta^+$) so, given a specified initial state $j(0)$, the rules (3) become

$$\begin{array}{l} j(\cdot) \text{ switches } 0 \curvearrowright 1 \text{ when } y(\cdot) \text{ increases across } \eta^+ \\ j(\cdot) \text{ switches } 1 \curvearrowright 0 \text{ when } y(\cdot) \text{ decreases across } \eta^- \end{array} \tag{4}$$

(again taking this as forbidding switching $0 \curvearrowright 1$ except when $y = \eta^+$ and similarly for $1 \curvearrowright 0$; a more complicated description would be needed if one wished to allow for jumps in $y(\cdot)$). Note that a nonideal relay is not expressible as a function: one cannot determine $j(t)$ from $y(t)$ when $\eta^- < y(t) < \eta^+$, although we do have a "Markovian" history dependence in that $j(t)$ is determined by only the most recent crossing. [The value of a nonideal relay for control-theoretic applications is that this separation might reduce the switching cost associated with 'chattering' [1] if the input would hover near the common threshold. Of course, it also occurs in "natural" contexts as the bacterial example.]

We can take $\mathbf{W}$ as characterized by the initial mode and the threshold pair (η^-, η^+), suppressing this in simply writing $\mathbf{W}[\cdot]$. [One could always work with the thresholds $0, 1$ by the input transformation $r \mapsto (r-\eta^-)/(\eta^+-\eta^-)$.]

One difficulty which arises in the use of (3) or (4) to define $\mathbf{W}$ is the possibility of a *Zeno phenomenon*: infinitely many modal switches within a finite time interval so there might not be a well-defined mode to switch from. However, we observe that

{1} *If the input function $y(\cdot)$ is continuous and the thresholds are separated as in* (4)*, then the Zeno phenomenon cannot occur.*

To see this, note that continuity of y on the compact interval $[0, T]$ implies uniform continuity so there is a minimum time needed for y to change from one threshold value to another. Thus, there is a bound on the number of interswitching intervals. [One easily sees that this bound is uniform for $y \in \mathcal{Y}$ with $\mathcal{Y}$ compact in $\mathcal{C}[0, T]$.]

Note that the switching formulation views the output transitions as instantaneous, ignoring the transient fine structures of the system dynamics and of the transducer itself as occurring on an unmodeled fast scale; this has possible consswquences which must affect our thinking, if not our subsequent analysis since we note that there is one anomalous situation in which (4) involves an ambiguity in the operation of $\mathbf{W}[\cdot]$: since (4) speaks only of $y(\cdot)$ *crossing* a threshold value, one must address the possibility that the input signal grazes this value without immediately crossing.

{2} *Consider a trajectory $y(\cdot)$ with $y(t_*) = \eta^+$ and $j(t_*-) = 1$. In this situation we generically expect that there would be neighboring trajectories $\hat{y}(\cdot)$ which do not merely graze and that these are of both types: some with $\hat{y} > \eta^+$ arbitrarily near t_* and some with $\hat{y}$ locally bounded away from η^+. For the former trajectories, we would switch $1 \curvearrowright 0$ near t_* and so might also expect to switch at t_* in the limit case. For the latter, however, we would continue past t_* without switching and so might also expect this in the limit.*

Certainly this is ambiguous. It is clearly possible to select unambiguously and, indeed, several of the treatments in [3], etc., select in such a way as to obtain various nice properties (as isotonicity, rate independence, etc.). Nevertheless, we will choose to let $\mathbf{W}$ be set-valued: given $y(\cdot) \in \mathcal{C}([0, T] \to H)$ we accept *all* modal functions satisfying the switching rules

$$\mathbf{W}[y(\cdot)] = \{j(\cdot) \in BV\,([0, T] \to \{0, 1\}) : (4)\,\} \tag{5}$$

The rationale for this definition (compare [6]) is that the model is to be viewed as a reduced description of a multiscale process so these possible discontinuities represent the effects of otherwise unmodeled behavior on a faster scale than is being considered. To the extent that there may be uncertainties or ambiguities involved in this reduced modeling, our response is to accept that any of these continuations is as consistent with our available information as any other; the decision to leave the fast time scale unmodeled precludes the rejection of either alternative. With this definition we now have the following.

{3} *When the index functions $\{j(\cdot)\}$ are topologized, e.g., by pointwise convergence, then* $\mathbf{W}$ *is a closed operator. Thus* $\mathbf{W}[\mathcal{Y}]$ *is compact whenever* $\mathcal{Y}$ *is compact in* $\mathcal{C}[0,T]$.

To see this, suppose we are given a sequence $\{j_n \in \mathbf{W}[y_n]\}$ in $\mathbf{W}[\mathcal{Y}]$ such that $y_n \to \bar{y}$ uniformly. By **{1}**, each j_n is then piecewise constant with a fixed number of switching times $t_{n,k}$ and is characterized by $\{t_{n,k}\}$ and the common interswitching values. By the compactness of $[0,T]$ one can then extract a further subsequence such that $t_{n,k} \to \bar{t}_k$ for each k. Since each j_n is piecewise constant with values in $\{0,1\}$, this gives pointwise convergence to a correspondingly piecewise constant $\bar{\jmath}$ and by (5) we have $\bar{\jmath} \in \mathbf{W}[\bar{y}] \subset \mathbf{W}[\mathcal{Y}]$.

We also observe that threshold determined switching can occur as a reduced model for certain singular perturbation problems and, conversely, those can be used to simulate a nonideal relay. Consider the system

$$\dot{x} = f(x,y) \qquad \varepsilon\dot{y} = g(x,y) \tag{6}$$

so the reduced system corresponds to solving $0 = g(x,y)$ to get $y = Y(x)$ (stable if $g_y < 0$) and then $\dot{x} = \varphi(x) = f(x, Y(x))$. However, if we consider $g(x,y) = x + 3y - y^3$, then there are stable branches $Y_-(x)$ for $x \leq 1$ and $Y_+(x)$ for $x \geq -1$. If, while following the upper branch $\dot{x} = \varphi_-$, one would reach $+1$ with $\varphi_-(1) > 0$, then (very rapidly for very small ε) y would increase to $Y_+(1)$ with little change in x. In the limit $\varepsilon \searrow 0$ this gives the instantaneous modal switching we have been describing; note that on taking ε very small this is just the multi-scale viewpoint. This is easy to visualize only when x, y are scalar variables, but the principle generalizes appropriately.

Having described operation only of the simplest nonideal relay (scalar input and $\{0,1\}$ output) so far in this section, we note that this is easily generalized. For example, we might consider $\mathbb{R}^d$-valued inputs and $\mathcal{M}$-valued outputs for finite $\mathcal{M}$. In this case we would define the operator $\mathbf{W}$ by switching rules in terms of specified sets $\mathcal{A}_{jk} \subset \mathbb{R}^d$ for $j \neq k \in \mathcal{M}$ with $j \neq k$ under the assumptions that

$$\begin{array}{ll} a) & \text{each } \mathcal{A}_{jk} \text{ is the closure of its interior} \\ b) & \text{for each } j, \text{ the sets } \mathcal{A}_{jk} \text{ are disjoint} \\ c) & \text{for each } k, \text{ the sets } \mathcal{A}_{jk},\ \mathcal{A}_{k\ell} \text{ are disjoint} \end{array} \tag{7}$$

The switching rules are then that:

$$j(\cdot) \text{ switches } j \curvearrowright k \text{ when } y(\cdot) \text{ enters } \mathcal{A}_{jk} \tag{8}$$

and, as earlier, we define the transducer by

$$\mathbf{W}[y(\cdot)] = \{j(\cdot) \in BV\,([0,T] \to \mathcal{M}) : (8)\,\} \tag{9}$$

{4} *This operator generalizes* (5)*, noting that* (4) *is of the form* (8) *with* $\mathcal{A}_{01} = [\eta^+, \infty)$ *and* $\mathcal{A}_{10} = (-\infty, \eta^-]$*. The behavior is now essentially the same with corresponding arguments. Given* (7)*, one has* $\{\mathbf{1}\}$*' when y is continuous. There is again the ambiguity as in* $\{\mathbf{2}\}$ *if y grazes* $\partial\mathcal{A}_{jk}$ *while one is in mode j and we accept appropriate choices to get* $\{\mathbf{3}\}$*'.*

As an example of such a modeling situation, one might think of a population of "agents" moving in Ω assigned to stations located at distinct points $\{\zeta_j\}$ with a switching rule that the assignment of an agent at $z \in \Omega$ would be switched $j \curvearrowright k$ if z became significantly closer to ζ_k than to ζ_j, e.g., if $|z - \zeta_k| < \frac{1}{2}|z - \zeta_j|$. [Whether this might satisfy ($7\,b$) would depend on the geometry of $\{\zeta_j\}$.] We note that the assumption ($7\,b$) is needed only to avoid another possible source of ambiguity — that a modal switch might be mandated with its destination indeterminate — and the assumption could be omitted if we would accept that.

This possibility cannot arise if there are only two modes and we will avoid further analysis of such complications by restricting our expository attention to the bimodal setting.

Another generalization might be to permit the thresholds to be time-dependent. If $\eta^{\pm}(t)$ are continuous, this does not change the analysis above, which could then instead be applied to $\hat{y} = y - \eta^+$, etc. [However, if $j(t-) = 0$ and $y - \eta^+$ jumped to above 0, then we would necessarily switch to mode 1; if it jumped from a negative value to 0, then one would have an anomalous point. Etc.]

3 A Lagrangian view of stochastic diffusion

We next consider the (random) motion of a single individual in a diffusing population from a Lagrangian point of view, now taking into account the coupling between the modal switching and the dynamics.

The simplest coupling (2) of the dynamics in (1) with a nonideal relay $\mathbf{W}$ as in Section 2 would take $v = v(z) = z$, giving the system

$$z' = \varphi_j(z) \qquad \text{with } j(t) \in \mathbf{W}[Y(z)](t) \tag{10}$$

(with initial conditions in the state space $\mathcal{Z} \times \mathcal{M}$ and with $Y : \mathcal{Z} \to \mathbb{R}$). Here, under fairly standard hypotheses, we could proceed by solving the dynamics ODE on the finite number of well-defined interswitching intervals.

[While one might think of $\mathcal{Z} = \mathbb{R}^d$, one has familiar examples involving infinite-dimensional $\mathcal{Z}$, e.g., diffusion with a global mode occurs for thermostat-controlled heating: $z \in \mathcal{Z} = \mathcal{C}(\Omega)$ is the temperature distribution, Y is point evaluation at the thermostat, and we are noting the threshold separation (bracketing the nominal setting) built into standard thermostats. Here the modes are given by PDEs: heat equations which differ by the effect of having the furnace ON or OFF.]

In this section we first consider a stochastic version of the simplest form of (10) — with z scalar and $Y(z) = z$. Our model now introduces some randomness into the dynamics and is to be modeled by the Ito equation

$$dz = \varphi_j(z)\, dt + \sigma_j\, dw \qquad \text{with } j(\cdot) \in \mathbf{W}[z(\cdot)] \tag{11}$$

where dw is the standard Brownian white noise. [For simplicity we are taking each σ_k to be constant. We are also restricting attention here to Brownian diffusion, but note the relevance of the developing theory of *anomalous diffusion*. Note that both $z(t)$ and $j(t)$ are random variables.] At this point we introduce the probability density

$$u^k(t,x)\, dx = Prob\,\{j(t) = k \text{ and } x \le z(t) \le x + dx\}, \tag{12}$$

conditioned on the initial probability distribution. Because of our switching rules, the partial density u^0 will be supported on the interval $\Omega_0 = (-\infty, \eta^+)$ while u^1 will be supported on $\Omega_1 = (\eta^-, \infty)$. The total probability is then

$$P = \int_{-\infty}^{\eta^+} u^0(x,t)\,dx + \int_{\eta^-}^{\infty} u^1(x,t)\,dx.$$

Of course, the initial data must be nonnegative with $P(0) = 1$.

It is immediate from the standard theory of Ito equations that the forward Kolmogorov equation (Fokker-Planck equation) for each will be a diffusion equation with drift, having the form

$$u_t = a_k\, u_{xx} - (\varphi_k(x)\, u)_x \qquad \text{with } a_k = \tfrac{1}{2}[\sigma_k]^2$$

for x away from the switching thresholds. Combining, we obtain a coupled system with a somewhat unusual domain, $\mathcal{Q} = (0,T) \times \Omega$, where Ω is the disjoint union of overlapping copies of the two intervals, Ω_0 and Ω_1. Thus,

$$\begin{aligned} {u^0}_t &= a_0\, {u^0}_{xx} - (\varphi_0(x)\, u^0)_x + \psi_0 \quad \text{for } x \in \Omega_0,\\ {u^1}_t &= a_1\, {u^1}_{xx} - (\varphi_1(x)\, u^1)_x + \psi_1 \quad \text{for } x \in \Omega_1. \end{aligned} \tag{13}$$

to which we adjoin the boundary conditions

$$u^0(\eta^+, t), \equiv 0 \qquad\qquad u^1(\eta^-, t) \equiv 0. \tag{14}$$

In (13), the source term ψ_0 is the contribution rate to u^0 (transition probability) of the modal switching $1 \curvearrowright 0$, occurring at the threshold $x = \eta^-$, and, complementarily, ψ_1 is the switching rate for $0 \curvearrowright 1$. To have $dP/dt = 0$ with these transitions localized at $\eta^\pm$, we easily see that ψ_0 must be a point source at η^- with rate $=$ $-$[outward flux of u^1]$=$ $a_1 {u^1}_x$, noting (14), and similarly for ψ_1. Thus, these sources must be

$$\psi_0 = a_1\, {u^1}_x\big|_{x=\eta^-}\ \delta(x - \eta^-), \qquad\qquad \psi_1 = -\, a_0\, {u^0}_x\big|_{x=\eta^+}\ \delta(x - \eta^+). \tag{15}$$

{5} *The system coupling* (13) *with* (14), (15) *has a unique solution for given L^2 initial data. If the initial data is nonnegative, so is the solution and the L^1 norm $P(\cdot)$ is preserved.*

To see this, we first obtain *a priori* estimates as usual on multiplying by u^0, u^1, noting that ${u^1}_x$ is smooth near η^- (away from the support at η^+ of its source term) so ψ_0 is bounded and similarly for ψ_1. Standard semigroup methods then give existence. Multiplying by the negative parts then gives nonnegativity, following [8] and just integrating then gives the preservation of P.

{6} A similar analysis would also apply to problems with time-dependent thresholds (always with $\eta^- < \eta^+$), although working with this is somewhat complicated by the fact that the parabolic domain $\mathcal{Q}$ (as described here) is no longer a cylinder. Although now applying at $\eta^\pm(t)$, the boundary conditions (14) are unchanged and, in view of this, the formula (15) (still obtained from $dP/dt = 0$) also remains unchanged — except that we must now think of each ψ_k as having two components: for transitions due to diffusion across a threshold (as above) and also for transitions due to motion of the threshold. [One must compute the flux in coordinates moving with the moving threshold so there is an additional term $u^1\,(d\eta^-/dt)$ but observe that the Dirichlet condition $u^1 = 0$ ensures that this addition vanishes. This will be important in Section 4 and will be further clarified there.]

To use more general $Y : \mathbb{R} \to \mathbb{R}$ for $\mathbf{W}[Y(z)]$ in (11), we need thresholds $Y^{-1}(\eta^\pm)$, expressed for z, and then consider (13) with

$$\Omega_0 = \{x : Y(x) < \eta^+\} \qquad\qquad \Omega_1 = \{x : Y(x) < \eta^-\}$$

with the boundary conditions: $u^k = 0$ on $\partial\Omega_k$. The sources ψ are given much as in (15) with the obvious modifications: e.g., ψ_0 is a sum of terms $\pm a_1 u^1_x\,\delta$ as in (15), now at $Y^{-1}(\eta^-)$ with $\pm$ signs taken as these are right or left endpoints of intervals of Ω_1. We will not consider time-dependent $Y = Y(\cdot, t)$ other than the nugatory modification above for moving thresholds, but do note that the geometry can become more complicated as the multiplicity of $Y^{-1}(\eta^\pm)$ may change.

Essentially the same holds for multidimensional diffusions (with z varying in $\Omega \subset \mathbb{R}^d$) using $y = Y(z, t)$; we do note the absence of a good notation for the measure (supported by $\partial\Omega_1$) which replaces $\delta(\cdot - \eta^-)$ in $(15)_0$ (with u^1_x replaced by $\nabla u^1 \cdot \mathbf{n}$), etc. We will actually be interested in settings where switching is determined by an auxiliary function v so we would take $y(t) = Y(v(t, z(t)))$

for an individual following a diffusive path $t \mapsto z(t) \in \Omega$. The regions Ω_0, Ω_1 above are now replaced by

$$\begin{aligned}\Omega_0 = \Omega_0(t) = \{x \in \Omega : Y(v(t,x)) < \eta^+\}\\ \Omega_1 = \Omega_1(t) = \{x \in \Omega : Y(v(t,x)) > \eta^-\}\end{aligned}$$

where each mode $0, 1$ would be permitted.

[For multimodal systems one could obtain comparable constructions of increasing complexity and, in support of our decision to restrict attention to the bimodal setting, we note that the descriptions here have already become decreasingly explicit with the increasing complexity of the constructions generalizing (13), (15).]

4 Switching modes in diffusing populations

We now turn to our principal concern: reaction/diffusion systems involving populations u of individuals which would follow stochastic individual trajectories and individually switch modes as in Section 2 when viewed from the Lagrangian perspective.

As already noted, we restrict ourselves for expository purposes to bimodal dynamics $\mathbf{u} = (u^0, u^1)$ given by a nonideal relay observing a coupled function v so, pointwise in x, we have $y(t,x) = Y(t, v(t,x))$ for the switching rules (4). In this section we will take y as given on $\mathcal{Q} = [0,T] \times \Omega$ and then in the next section will consider the problem with a coupled sysyem for the combined u, v. [For expository simplicity we take v scalar and $Y(r) = r$ so there is no distinction between y and v. We do note that our formulation would handle a vectorial auxiliary function v (e.g., as seen in [9, XI.8 Figure 4.]) without significant change.]

The systems under consideration are somewhat similar to (13) except that u^0, u^1 are no longer probabilities, but concentration densities for a diffusing population which now may also involve interaction/growth and so need no longer be precisely conservative. Formally this can be written as

$$\begin{aligned} {u^0}_t - a_0 \Delta u^0 \ &= f_0(u^0, u^1, \, v) + \psi_0 - \psi_1 \\ {u^1}_t - a_1 \Delta u^1 \ &= f_1(u^0, u^1, \, v) + \psi_1 - \psi_0 \end{aligned} \tag{16}$$

where the terms $\psi_{0,1}$ denote the respective rates of transitions $1 \curvearrowright 0$ contributing to u^0 and of $0 \curvearrowright 1$ contributing to u^1. We assume throughout that Ω is bounded with a sufficiently smooth boundary $\partial\Omega$ and that the interaction terms satisfy

$$\begin{aligned} &f_0, f_1(u^0, u^1, v) \geq 0 \quad \text{if either } u^k \leq 0, \\ &f(u^0, u^1, v) = f_0 + f_1 \leq C(1 + u^0 + u^1). \end{aligned} \tag{17}$$

Our principal task at this point is to model $\psi_{0,1}$ along the lines of (15) so that the problem becomes meaningful. We begin with the single key assumption that y is continuous on the compact set $\mathcal{Q}$ — which ensures that $\mathcal{B}_0$, $\mathcal{B}_1$ (and *a fortiori* $\mathcal{A}_0\,\mathcal{A}_1$) are separated by a distance depending only on $[\eta^+ - \eta^-]$ and the uniform modulus of continuity of y. It is then convenient to partition $\mathcal{Q}$ into regions

$$\begin{aligned} &\mathcal{A}_0 = \{(t,x) \in \mathcal{Q} : y(t,x) < \eta^-\}, \\ &\mathcal{A}_1 = \{(t,x) \in \mathcal{Q} : y(t,x) > \eta^+\}, \\ &\mathcal{A}_* = \{(t,x) \in \mathcal{Q} : \eta^- < y(t,x) < \eta^+\} \\ \text{and} \qquad &\mathcal{B}_0 = \{(t,x) \in \mathcal{Q} : y(t,x) = \eta^-\}, \\ &\mathcal{B}_1 = \{(t,x) \in \mathcal{Q} : y(t,x) = \eta^+\} \end{aligned} \tag{18}$$

with the obvious interpretation of $\mathcal{A}_0(t)$, etc., noting the distinction between $\partial[\mathcal{A}_0(t)]$ and $[\partial\mathcal{A}_0](t)$, etc. In terms of this notation it is now complicated but relatively straightforward to describe the model we want for modal transition $1 \curvearrowright 0$ and similarly for $0 \curvearrowright 1$.

Using a Lagrangian viewpoint to follow an individual, we see that the mode of any individual of the population u^1 must change $1 \curvearrowright 0$ (going forward in t, which we visualize as going upward in $\mathcal{Q}$) if she moves into the region $\mathcal{A}_0$ — with an allowed possibility of change while in $\partial\mathcal{A}_0$. At this point the geometry of the partition (18) becomes quite important and we begin our attempt to visualize this under the simplifying assumption that $\mathcal{B}_{0,1}$ are nice surfaces in $\mathcal{Q}$.

Note that the support of ψ_0 must lie in the underside and lateral surface of $\partial\mathcal{A}_0$ where such particles might be entering from $\mathcal{A}_*$ with increasing t. We are, of course, imposing the condition that $u^1 \equiv 0$ on $\mathcal{A}_0$ so, anticipating some regularity in x across $\partial[\mathcal{A}_0(t)]$, this becomes a homogeneous Dirichlet condition at $\partial[\mathcal{A}_0(t)]$ for u^1 on $\mathcal{A}_*$. Thus, as already noted in $\{\mathbf{6}\}$, there is no contribution from boundary motion in using the flux (from $\mathcal{A}_*$ through $\mathcal{B}_0$ toward $\mathcal{A}_0$ to become part of u^0) as the relevant contribution to ψ_0.

{7} Unfortunately, a technical complication arises: this analysis fails to complete the story if some part of the underside of $\mathcal{A}_0$

might be flat with $[\partial\mathcal{A}_0](t)$ having nonempty interior ω. In this case u^1 need not be 0 there: one might have a nonvanishing concentration of mode 1 there at $t-$ in $\mathcal{A}_*$, coming into ω and so into $\mathcal{A}_0$ from $t-$ to $t+$. In this situation, where $\omega \not\subset \partial[\mathcal{A}_*(t-)]$ there is then an instantaneous mode transition $1 \curvearrowright 0$ for all the individuals comprising $\{u^1(t-,x) : x \in \omega\}$ to become part of $\{u^0(t+,x) : x \in \omega\}$; should this occur, we denote this contribution to ψ_0 by $\tilde{\psi}_0$, etc.

Note that if we have such a flat place and were to adjust v to tilt this slightly, then we would have $\tilde{\psi}_0$ obtained by the flux relative to a rapidly moving spatial interface between $[\partial\mathcal{B}_0](t)$ and $[\partial\mathcal{A}_0](t)$ with u^1 then vanishing at this interface. To the extent that this is a good approximation, we would expect that u^1 must at least be small there if not exactly 0. Nevertheless, there are some technical difficulties handling $\tilde{\psi}_0$ directly if there could be infinitely many instants t at which such flat places would occur, making a difficulty comparable to the Zeno phenomenon of Section 2. In this setting the separation of regions coming from the continuity assumption may be insufficient to avoid this and for present purposes we could simply require by hypothesis that v is such that these do not occur at all (e.g., having adjusting so each would be tilted) whence $\tilde{\psi}_0 = 0$ and similarly $\tilde{\psi}_1 = 0$. Instead, however, we will assume the diffusion coefficients are the same, enabling us to handle the Zeno phenomenon fairly easily.

The treatment of $\tilde{\psi}_k$ requires some care with respect to regularity in t and the formulation of the model: we must have the requisite jumps where the flat places occur, giving a $BV[0,T]$ component of the solution pointwise on Ω and then, excluding those jumps, a weak form

$$\begin{aligned}
&\int_\Omega [\varphi_0\, {u^0}_t + a\nabla\varphi_0 \cdot \nabla u^0] \\
&\qquad = \int_\Omega \varphi\, f_0 - \int_{\partial[\mathcal{A}_0(t)]} \varphi_0\, [a\nabla u^1 \cdot \mathbf{n}] + \int_{\partial[\mathcal{A}_1(t)]} \varphi_0\, [a\nabla u^0 \cdot \mathbf{n}] \\
&\int_\Omega [\varphi_1\, {u^1}_t + a\nabla\varphi_1 \cdot \nabla u^1] \\
&\qquad = \int_\Omega \varphi\, f_1 - \int_{\partial[\mathcal{A}_1(t)]} \varphi_1\, [a\nabla u^0 \cdot \mathbf{n}] + \int_{\partial[\mathcal{A}_0(t)]} \varphi_1\, [a\nabla u^1 \cdot \mathbf{n}]
\end{aligned} \tag{19}$$

holding for smooth φ_0, φ_1 satisfying Neumann boundary conditions at $\partial\Omega$.

{8} *The selection of an admissible space for solutions includes the specifications that $u^1 \equiv 0$ on the set $\mathcal{A}_0$ defined by the given v and that $u^0 \equiv 0$ on $\mathcal{A}_1$ as well as homogeneous Neumann boundary conditions (no flux) for u^0 and for u^1 at $\partial\Omega$. For given continuous v and such suitable data,* (19) *has a suitable solution.*

The proof of {**8**} is based on construction of a suitable finite element approach to approximation (whose description we defer to the next section) and then a compactness argument based on a sequence of a priori estimates which we sketch here.

We begin by showing the nonnegativity of solutions. The treatment of $\tilde{\psi}_0$ cannot make $u^1(t+, x)$ negative as $(t+, x) \in \mathcal{A}_0$ where $u^1 \equiv 0$. With $u^1 \geq 0$, we see that the outward flux must be nonpositive where we are imposing on it a homogeneous Dirichlet condition so the contribution ψ_0 must be nonnegative. Completion of this argument is then standard, from the first condition in (17).

With equal diffusion coefficients and consistent bounded nonnegative initial data we will have bounded (L^∞) solutions. To see this, set $u = u^0 + u^1$ and $f = f_0 + f_1$. Then add the equations of (19) with $\varphi_0 = \varphi_1 = \varphi$ to get

$$\int_\Omega [\varphi u_t + a\nabla\varphi \cdot \nabla u] = \int_\Omega \varphi f$$

since the ψ terms cancel. Now use $g(t)$ satisfying $\dot{g} = C(1+g)$ with $g(0) \geq \max\{u(0,\cdot)\}$ to take $\varphi = [u-g]_+ = \max\{u-g,\, 0\}$, noting that where $\varphi \neq 0$ one has

$$\begin{aligned} &0 < \varphi = u - g, \qquad f_t = u_t - \dot{g}, \qquad \nabla\varphi = \nabla u \\ \text{so} \quad &C(1+u) = \dot{g} + C\varphi, \quad \varphi u_t = [\tfrac{1}{2}\varphi^2]_t + \dot{g}\,\varphi, \quad \nabla\varphi\cdot\nabla u = |\nabla\varphi|^2. \end{aligned}$$

Since φ vanishes at $t = 0$, Gronwall's Inequality then gives $\varphi \equiv 0$ so $0 \leq u^0, u^1 \leq u \leq g$ pointwise. [Even with unequal diffusion coefficients one could have easily bounded u in $L^1(\Omega)$, but we will need the L^∞ estimate to handle the possibility of a Zeno difficulty from $\tilde{\psi}$.]

Using $\varphi_0 = u^0$ in (19) and integrating over $[0,t]$ gives

$$
\begin{aligned}
&\tfrac{1}{2}\|u^0(t,\cdot)\|^2 + \int_0^t a\|\nabla u^0\|^2 \\
&\quad = \tfrac{1}{2}\|u^0(0,\cdot)\|^2 + \int_0^t \langle u^0, \varphi_0\rangle + \int_0^t \int_{\partial[\mathcal{A}_0(t)]} u^0\,[-a\nabla u^1\cdot\mathbf{n}] + \sum_0^t \langle u^0, \tilde{\psi}_0\rangle
\end{aligned}
$$

and we must estimate the last two terms on the right.

The final term is, apparently, an uncountable sum, although we have pointwise bounds both for u_0 and for $\tilde{\psi}_0$. The separation of $\mathcal{A}_0, \mathcal{A}_1$, however, ensures a fixed finite bound (uniform in $x \in \Omega$) on the number of nonzero summands occurring for that x, giving a bound for $\int_\Omega \sum_0^t u^0\psi_0$.

To estimate the penultimate term we note that the same separation permits us to cover the timelike portion of $\partial\mathcal{A}_0$ by a finite set of cylinders (with size bounded below) in $\mathcal{A}_*$ on which ψ_1 does not occur since this is supported on $\partial\mathcal{A}_1$. For each of these we use a cutoff function χ (with $|\nabla\chi|$ uniformly bounded) to restrict attention to that cylinder with homogeneous Dirichlet data and then estimate the relevant outward flux ψ_0 integrated over that part of $\partial\mathcal{A}_0$; since we already have a bound on u^0, summing these contributions gives the desired estimate.

Thus we obtain the usual $L^2(\mathcal{Q})$ regularity for ∇u^0 and similarly for ∇u^1. This analysis also gives a BV estimate for the variation in t. We emphasize that all these estimates are uniform for auxiliary functions y in any compact subset of $C(\mathcal{Q})$, whence uniform separation. This provides the compactness needed to get a convergent sequence of finite element approximations and so completes the argument for existence $\{\mathbf{8}\}$.

5 The coupled problem

In this section we arrive at our final result: the coupling of the pair of equations (16) for u^k (taken with $a_0 = a_1 = a$) with a similar diffusion equation for the auxiliary function v:

$$
\begin{aligned}
{u^0}_t - a\Delta u^0 &= f_0(u^0, u^1, v) + \psi_0 - \psi_1 \\
{u^1}_t - a\Delta u^1 &= f_1(u^0, u^1, v) + \psi_1 - \psi_0 \\
v_t - a_*\Delta v &= f_*(u^0, u^1, v)
\end{aligned}
\tag{20}
$$

with the Neumann boundary conditions

$$-a\nabla u^k \cdot \mathbf{n} = 0, \qquad -a_*\nabla v \cdot \mathbf{n} = \sigma \qquad\qquad \text{on } \partial\Omega \tag{21}$$

and with suitable initial conditions, asking that $u^k(0,\cdot)$ be nonnegative and pointwise bounded and consistent with the continuous function $v(0,\cdot)$ on Ω.

For our purposes, it will be crucial that the function v and so $y = Y(v)$ should be continuous so the u^k equations are to be interpreted as in **{8}** of Section 4 with a minimum width $d_* > 0$ separating $\mathcal{A}_0 \cup \mathcal{B}_0$ from $\mathcal{A}_1 \cup \mathcal{B}_1$.

{9} Our strategy will be to introduce spaces $\mathcal{V}^h$ of piecewise affine functions on triangulations Ω^h of Ω and projections $\mathbb{P}^h$ onto the spaces $\mathcal{A}^h$ of piecewise constants (e.g., by averaging or by evaluation at chosen points). If these are regular triangulations with mesh width h, we may consider letting $h \to 0+$ and will have $\mathcal{V}^h$ becoming dense in $C(\mathcal{Q})$ and also, for continuous v, will have $y^h = Y(\mathbb{P}^h v) \to y = Y(v)$ — uniformly for v in any compact subset of $C(\mathcal{Q})$. Further, given $\mathcal{V}^h, \mathbb{P}^h$ we may define the partition $\mathcal{A}_0^h$, etc., as in (18), by

$$\mathcal{A}_0^h = \{(t,x) : y(\mathbb{P}^h v) < \eta^-\},$$

etc., noting that each such set is a union of a finite number of sets of the form $(t,t') \times \omega$ where $\omega \in \Omega_h$. The finiteness of the t-partition comes from **{1}** and is uniform over h (provided h is small compared with the width d_*) and over the compact set for v: the separation, in time as well as space, of $\mathcal{A}_0^h$ from $\mathcal{A}_1^h$ will then be uniform.

We now want to use the weak form, corresponding to (19), of the u^k equations; it is convenient to leave the v equation as in (20), (21). We refer to this partially discretized system as $(20)^h$ although we have, so far, only described it between the modal switching times for y^h. We do note that, as is usual, $(20)^h$ is just an ODE between these switching times with $\mathcal{A}_0^h$, etc., here independent of t: the modal switching of individual particles occurs only as they cross the threshold boundaries. Note that the contributions $\psi_{0,1}$ here represent gradient discontinuities at those faces of the triangulation corresponding to an interior Dirichlet condition imposed by having $u^1 \equiv 0$ on an element belonging to $\mathcal{A}_0^h$. Etc. The transition points of $(20)^h$ correspond to times when some component of $\mathbb{P}^h v$ hits an implicit threshold $y^{-1}(\eta^\pm)$. We would be solving the v equation forward in t along with this ODE and, while we will see that v is continuous, note that $\mathcal{A}_0^h$ may change from $t-$ to $t+$: if so with the element $\omega \in \mathcal{A}_*^h(t-)$ but $\omega \in \mathcal{A}_0^h(t+)$, then $u^1(t-)$ is added to u^0 for that ω as (part of) $\tilde{\psi}_0$ and the system is then restarted. [Note that until we restart and continue the evolution of v we do not know definitively about such a change of $\mathcal{A}_0^h$, so this may become an anomalous point of the discretized system, treated as in **{2}**.] It is easy to see that $u^{k,h} \geq 0$ by this construction and (modifying the second condition of (17) to: $f, |f_*| \leq C(1 + u^0 + u^1 + |v|)$ now) we have, much as in Section 4, a pointwise bound on $u^{k,h}$ depending only on the constant C of the modified (17) and a bound on the initial data, but not on h.

The v equation in $(20)^h$ is quite standard so there is no difficulty with solution of that. With the growth condition on f_* (and the resulting bound for $u^{k,h}$) one similarly has a sup norm bound on f_* as it appears there so, with σ fixed (or even just suitably bounded), one can get $v = v^h$ in a compact subset of $C(\mathcal{Q})$ whence, as in **{3}**, we have a uniform bound on the number of switching times for $(20)^h$.

With the estimates noted, this ensures convergence in $C(\mathcal{Q})$ of $u^{k,h} \to u^k$ and $v^h \to v$ for some subsequence. Since we have obtained a bound on $\psi_{0,1}$ in the dual space $[C(\Omega)]^*$, we have subsequential convergence for that, with the correct limit by consistency; a similar result holds for $\tilde{\psi}_k$. It then follows that such a subsequential limit is a solution of (20). We have thus proved our desired result:

Theorem 5.1. *Let Ω be a bounded region in $\mathbb{R}^d$ with Lipschitzian boundary $\partial\Omega$ and let $y : \mathbb{R} \to \mathbb{R}$ be continuous and $\eta^- < \eta^+$. We then consider* (20) *in the sense described above, as corresponding to modal transitions determined by $y = Y(v)$ by the nonideal relay as $\mathbf{W}[y(t, \cdot)]$ (with thresholds $\eta^\pm$) for the motion of each diffusing particle.*

Suppose we are given (21) *and bounded initial data consistent with this and with $v(0, \cdot)$ continuous. Assume further consistency in requiring, at $t0$, that $u^1 = 0$ where $y(v) \leq \eta^-$ and that $u^0 = 0$ where $y(v) \geq \eta^+$. Then there is a suitable solution of the problem* (20)

References

[1] I. Capuzzo-Dolcetta and L.C. Evans, *Optimal switching for ordinary differential equations*, SIAM J. Cont. Opt., **22**, pp. 143–161 (1984).

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

[3] M.A. Krasnosel'skĭi and A.V. Pokrovskiĭ, *Systems with Hysteresis*, Nauka, Moscow, 1983 [*transl.,* Springer-Verlag, Berlin, 1989].

[4] S. Lenhart, T.I. Seidman, and J. Yong, *Optimal control of a bioreactor with modal switching*, Math. Models and Methods in Appl. Sci. **11**, pp. 933–949, (2001).

[5] T.I. Seidman, *Switching systems: thermostats and periodicity,* Math. Research Report **83-07**, UMBC, 1983. http://userpages.umbc.edu/∼seidman/ss_83.pdf

[6] T.I. Seidman, *Some aspects of modeling with discontinuities,* Int'l. J. Evolution Eqns. **3**, pp. 129–143, (2008).

[7] T.I. Seidman, *Optimal control of a diffusion/reaction/switching system,* DCDS-B, to appear, (2013).

[8] G. Stampacchia, *Equations elliptiques du second ordre á coefficients discontinues,* Les Presses de l'Université de Montréal, 1966.

[9] A. Visintin, *Differential Models of Hysteresis,* Springer-Verlag, Berlin, 1994.