The error in nonsmooth dynamics

Mike R. Jeffrey
Engineering Mathematics, University of Bristol, UK, email: mike.jeffrey@bristol.ac.uk

The method of solving piecewise-smooth differential equations by interpolating across their discontinuities is discussed. An example is given of stick-slip motion of a block on a surface, where the discontinuity of Coulomb's friction force is resolved using Filippov's standard method of differential inclusions. We first show that this method cannot be justified in general, because the equations either side of a discontinuity do not restrict dynamics *on* the discontinuity to the extent that Filippov's method assumes. We then show that allowing for unknown errors in the model restores Filippov's dynamics. The outcome is that a balance between accuracy and precision of a model determine when piecewise-smooth systems methods are applicable. The result is particularly important in understanding how discontinuities relate to singular limits of smooth systems. In a practical sense the effect of error here is analagous to the practical effect of dither: the presence of an imperfection to eliminate the appearance of undesirable structure, in this case, lack of uniqueness at a discontinuity.

One assumes that underlying any system are a set of well-determined, and more-or-less smoothly varying, physical laws. Nevertheless, smooth variations can give rise to discontinuities by means of, for example, bifurcations, shocks, or singular perturbations. If an observer is able to reconstruct a set of physical laws only at the piecewise-smooth level, i.e. to the extent that they involve a discontinuity like an impact or a switch, can the system dynamics be uniquely determined? In other words, does determinism survive a discontinuity?

Discontinuities arise in the mechanics of stick-slip dynamics or impact. They describe control switches in electronics and in living nervous systems, and transitions in social, political, or economic models; see e.g. [7]. Their practical utility suggests that differential equations can have physically sensible solutions even at a discontinuity.

The theory of *piecewise-smooth* dynamics hails from the Russian literature [1, 5, 14], from which a very idealized notion of a 'solution' has grown. It is becoming apparent that this traditional view is overly restrictive (see e.g. [6]). The aim here is to lift the veil slightly by giving formal expression to an irony, that the idealized solution in common use is justified not by ignoring imperfections in the discontinuous model, but by carefully accounting for their perturbative effect.

Consider the model of a system that switches sharply between different regimes of behaviour, say $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}; c)$ where the parameter c switches between values c^+ and c^- when some function $h(\mathbf{x})$ changes sign. In the case of friction, for example, $c = \mu F \text{sign}(h)$ might be the force between two rough surfaces with relative speed $h = u$, for some normal force F and friction coefficient μ. Typically we have no detailed knowledge of the equations that apply at $h = 0$ as the switch occurs.

We have then a prototype discontinuous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}; c) = \begin{cases} \mathbf{f}(\mathbf{x}; c^+) & \text{if} \quad h(\mathbf{x}) > 0 \ , \\ \mathbf{f}(\mathbf{x}; c^-) & \text{if} \quad h(\mathbf{x}) < 0 \ , \end{cases} \tag{1}$$

where $\mathbf{x}$ and $\mathbf{f}$ are vectors. The righthand side is not defined when $h = 0$, reflecting a lack of knowledge of (possibly complex) processes that occur during switching between $\mathbf{f}(\mathbf{x}; c^+)$ and $\mathbf{f}(\mathbf{x}; c^-)$ as h changes sign.

In practice we may know only the function values

$$\mathbf{f}^+(\mathbf{x}) = \mathbf{f}(\mathbf{x}; c^+) \quad \& \quad \mathbf{f}^-(\mathbf{x}) = \mathbf{f}(\mathbf{x}; c^-) \ , \tag{2}$$

without explicitly knowing the c-dependence of $\mathbf{f}(\mathbf{x}; c)$. With this in mind, Filippov completed the system (1) by interpolating between $\mathbf{f}^+$ and $\mathbf{f}^-$ at $h = 0$, making a differential inclusion. There is no unique way to do this, but Filippov (and the majority since) focussed on the convex combination,

$$\dot{\mathbf{x}} = \lambda \mathbf{f}^+(\mathbf{x}) + (1 - \lambda)\, \mathbf{f}^-(\mathbf{x}) \ , \qquad \lambda \in [0, 1] \ . \tag{3}$$

This physically reasonable result arises, for example, in the limit of small hysteresis or noise in the switching process (see e.g. [4–6, 9, 14]). An interpretation which we will find useful is to consider λ as the fraction of any time instant for which $\dot{\mathbf{x}}$ is dominated by $\mathbf{f}^+$, and $1 - \lambda$ as the remaining fraction for which $\dot{\mathbf{x}}$ is dominated by $\mathbf{f}^-$.

As stated above, this is a reasonable convention, but a convention nonetheless. The hope that Filippov's differential inclusion is the correct choice has given birth to a thriving piecewise-smooth dynamical systems community, of broad use from designing electronic controllers to studying population dynamics. Unfortunately, there are infinitely many alternatives to (3) that do not give equivalent dynamical solutions. Moreover there exist situations where (3) is demonstrably incorrect, seeming to condemn this approach to be esoteric and system-specific.

This letter is devoted to analysing if, and when, Filippov's method is valid. We first show that the method is not generally correct if we assume a model to be exact, and this has implications, in particular, for attempts to approximate the discontinuity by a smooth transition. However, since any model is only an idealization that ignores certain errors, we can isolate the only robust part of the dynamics, which we show is indeed the commonly used result, namely (3). Moreover, conditions can be found for when this rule applies.

Let us assume that the switch actually takes place within a small neighbourhood $|h| < \varepsilon$ for some small constant ε, so (1) becomes

$$\dot{\mathbf{x}} = \begin{cases} \mathbf{f}^+(\mathbf{x}) & \text{if} \quad h(\mathbf{x}) > +\varepsilon \ , \\ \mathbf{f}^-(\mathbf{x}) & \text{if} \quad h(\mathbf{x}) < -\varepsilon \ . \end{cases} \tag{4}$$

We can set $\varepsilon = 0$ at any time to obtain the original discontinuous system, but resisting this affords a more general argument. To continue the system to $|h| \le \varepsilon$, let (4) be subsumed into

$$\dot{\mathbf{x}} = f\left(x;\ \lambda(h)c^+ + [1-\lambda(h)]\,c^-\ \right)\ , \tag{5}$$

introducing a function $\lambda(h)$ that satisfies

$$\lambda(h) \in \begin{cases} 1 & \text{if } \ h > +\varepsilon\ , \\ 0 & \text{if } \ h < -\varepsilon\ , \\ [0,1] & \text{if } \ h \le \varepsilon\ , \end{cases} \tag{6}$$

so that the righthand side equals (4) for $|h| > \varepsilon$. Equation (5) is an alternative to (3) which interpolates between the parameter values $c^\pm$ instead of the function values $\mathbf{f}^\pm$ inside $|h| \le \varepsilon$. Are the two consistent? An attempt to write (5) closer to the form (3) gives

$$\dot{\mathbf{x}} = \phi(\mathbf{x})\mathbf{f}^+(\mathbf{x}) + (1-\phi(\mathbf{x}))\,\mathbf{f}^-(\mathbf{x})\ , \tag{7}$$

using ϕ in place of λ. Unlike (3), however, where λ was limited to the interval $[0,1]$, the function ϕ is given by

$$\phi(\mathbf{x}) = \frac{[\mathbf{f}(\mathbf{x};c) - \mathbf{f}^-(\mathbf{x})]\cdot\nabla h(\mathbf{x})}{[\mathbf{f}^+(\mathbf{x}) - \mathbf{f}^-(\mathbf{x})]\cdot\nabla h(\mathbf{x})}\ , \tag{8}$$

which depends on $\mathbf{x}$, not just on $h(\mathbf{x})$. So if $\mathbf{f}(\mathbf{x};c)$ is unknown for $|h| \le \varepsilon$, then the way ϕ varies with $\mathbf{x}$ near $h = 0$ is unknown. To make this lack of knowledge more explicit let us seek a form closer to (3), insisting that the vector fields are multiplied by functions of h only. This is achieved by introducing a new function $\mu(h)$ that satisfies

$$\mu(h) \in \begin{cases} 0 & \text{if } \ |h| > \varepsilon\ , \\ [0,1] & \text{if } \ |h| \le \varepsilon\ , \end{cases} \tag{9}$$

and writing

$$\begin{aligned}\dot{\mathbf{x}} = &\left[\lambda(h)\mathbf{f}^+(\mathbf{x}) + (1-\lambda(h))\,\mathbf{f}^-(\mathbf{x})\right](1-\mu(h)) \\ &+\ \mu(h)\mathbf{f}^\varepsilon(\mathbf{x})\,,\end{aligned} \tag{10}$$

by subtituting $\mathbf{f}^\varepsilon = \mathbf{f}^- + (\mathbf{f}^+ - \mathbf{f}^-)\,(\lambda\mu - \lambda + \phi)\,/\mu$. For some functions λ and μ subject to (6) and (9), and a function $\mathbf{f}^\varepsilon$ which is unknown (since $\phi(\mathbf{x})$ is unknown in $|h| < \varepsilon$), the system (10) is actually the *general way of regularizing* the problem (1) across the discontinuity.

Filippov's convention (3) assumes that μ is identically zero, but we have no justification for this in $|h| < \varepsilon$. To see why, we can interpret the quantities λ, $1-\lambda-\mu$, and μ, as fractions of any time instant for which the dynamics is dominated by each vector field $\mathbf{f}^+$, $\mathbf{f}^-$, and $\mathbf{f}^\varepsilon$, respectively; this can be called the *residence time* in each field (following [10]). Since trajectories can slide along the discontinuity for arbitrarily long times, the fraction μ may be close to unity, and $\mu\mathbf{f}^\varepsilon$ may not be negligible.

The outcome of this is that the velocity $\dot{\mathbf{x}}$ depends on $\mathbf{f}^\varepsilon$ (or $\mathbf{f}^0$ if $\varepsilon = 0$). This is fatal because the function $\mathbf{f}^\varepsilon$ is unknown and need not be restricted by the values $\mathbf{f}^+$ or $\mathbf{f}^-$. The $\mathbf{f}^\varepsilon$-dependent dynamics near the discontinuity is therefore entirely unknown. Filippov himself comments on this [5], but leaves the notion that his method works only for systems 'close to' the $\mu = 0$ model, in a sense that has never fully been established.

The outcome from (10) is that in general physical applications there is no way to uniquely "fill in" a discontinuity in a differential equation. Since the success of discontinuous dynamics in applications suggests otherwise, let us reconsider why we write a model in a form like (1) to begin with, then see where the result (10) leads us.

When modeling even simple real world systems we neglect many sources of error, some knowingly, and some which can never be quantified. We rely on the stability of a system to preserve its qualitative behaviour despite such errors, and we are wary that nonlinearity can make this fail spectacularly, as happens with chaos. In the problem above it seems that the system becomes uncertain near the discontinuity, particularly when trajectories spend a significant time in its vicinity. The source of uncertainty is our lack of knowledge of the exact laws of some real system, so let us consider how effects omitted from our model would affect solutions of (10).

If changes in state over small time intervals δt take the form $\delta\mathbf{x}(\mathbf{x}) = \mathbf{f}(\mathbf{x})\delta t$, we write an empirical model for the system as $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. Then let us assume that this contains errors in both the state and the vector field, so that the true system evolves according to $\delta\mathbf{x} = \mathbf{f}(\mathbf{x} + \kappa\boldsymbol{\xi}(\mathbf{x})) + \kappa\mathbf{g}(\mathbf{x})$, where κ is a small constant, while $\boldsymbol{\xi}$ and $\mathbf{g}$ are finite vector functions (if these are time-dependent we can include t in the vector $\mathbf{x}$). The shifts $\boldsymbol{\xi}$ and $\mathbf{g}$ may account for effects such as hysteresis, delay, random noise, or other neglected errors; the important point is that they are not known exactly.

When we attempt to integrate small intervals $\delta\mathbf{x}$ to find a solution $\mathbf{x}(t) = \int \delta\mathbf{x}(t)$, we have

$$\mathbf{x}(t) = \int_0^t dt\mathbf{f}(\mathbf{x}(t) + \kappa\boldsymbol{\xi}(\mathbf{x}(t))) + \kappa\int_0^t dt\mathbf{g}(\mathbf{x}(t))\ . \tag{11}$$

Expanding $\mathbf{f}$ in powers of κ we can write $\mathbf{f}(\mathbf{x} + \kappa\boldsymbol{\xi}) = f^i(\mathbf{x}) + \kappa\boldsymbol{\xi}f'(\mathbf{x}) + \mathsf{O}\left(\kappa^2\right)$, where i is either $+$, $-$, or ε depending on the value of $h(\mathbf{x} + \kappa\boldsymbol{\xi})$. We therefore have

$$\begin{aligned}\mathbf{x}(t) = &\int_0^t dt\left\{\begin{array}{ll}\mathbf{f}^\pm(\mathbf{x}) & \text{if } \ h(\mathbf{x}+\kappa\boldsymbol{\xi}(\mathbf{x})) > \varepsilon \\ \mathbf{f}^\varepsilon(\mathbf{x}) & \text{if } \ h(\mathbf{x}+\kappa\boldsymbol{\xi}(\mathbf{x})) < \varepsilon\end{array}\right\} \\ &+\kappa\int_0^t dt\boldsymbol{\xi}(\mathbf{x}(t)) + \kappa\int_0^t dt\mathbf{g}(\mathbf{x}(t)) + \mathsf{O}\left(\kappa^2\right).\end{aligned} \tag{12}$$

The second line is of order κ (though for large time periods, of order $1/\kappa$, these integrals can give significant deviations, as in chaotic systems for example). So the integral is given, for moderate time periods at least, by

$$\mathbf{x}(t) = \int_0^t dt\left\{\begin{array}{ll}\mathbf{f}^\pm(\mathbf{x}) & \text{if } \ h(\mathbf{x}+\kappa\boldsymbol{\xi}(\mathbf{x})) > \varepsilon \\ \mathbf{f}^\varepsilon(\mathbf{x}) & \text{if } \ h(\mathbf{x}+\kappa\boldsymbol{\xi}(\mathbf{x})) \le \varepsilon\end{array}\right\} + \mathsf{O}\left(\kappa\right)\ . \tag{13}$$

We can assign weights to these two outcomes. For simplicity assume $|\boldsymbol{\xi}| = 1$, and express the probability that $\mathbf{x} + \kappa\boldsymbol{\xi}$ lies inside the ε-neighbourhood of $h = 0$, as

$$P[h(\mathbf{x} + \kappa\boldsymbol{\xi}(\mathbf{x})) \le \varepsilon] = \frac{\int_{S_\kappa(\mathbf{x})\cap|h|<\varepsilon} a(\mathbf{x})dS}{\int_{S_\kappa(\mathbf{x})} a(\mathbf{x})dS}\ , \tag{14}$$

integrating over surface elements dS on the (n dimensional) spherical shell $S_\kappa(\mathbf{x})$ of radius κ centered on $\mathbf{x}$.

The function $a(\mathbf{x})$ gives a probability distribution for $\boldsymbol{\xi}$. If $\mathbf{x}$ is uniformly distributed on the unit circle we have

$$P[h(\mathbf{x}+\kappa\boldsymbol{\xi}(\mathbf{x}))\le\varepsilon] = \frac{\text{area of } S_\kappa(\mathbf{x})\cap\{\mathbf{x}:|h(\mathbf{x})|<\varepsilon\}}{\text{surface area of } S_\kappa(\mathbf{x})} \le \frac{\varepsilon \text{ band around equator of } S_\kappa(\mathbf{x})}{\text{surface area of } S_\kappa(\mathbf{x})} \approx \frac{4\varepsilon}{2\pi\kappa} = \frac{2\varepsilon}{\pi\kappa}$$

in two dimensions, and with different expressions in $n>2$ dimensions, but always of order ε/κ. If the distribution is not uniform, $\boldsymbol{\xi}$ could have a preference to lie along $h=0$ which will raise the probability that the corrected location $\mathbf{x}+\kappa\boldsymbol{\xi}$ lies in $|h|<\varepsilon$. Consider instead a Gaussian distribution $a=e^{-h^2}$, so the error tends to push $\mathbf{x}$ along the discontinuity. Again, in two dimensions but easily generalized, we have,

$$P[h(\mathbf{x}+\kappa\boldsymbol{\xi}(\mathbf{x}))\le\varepsilon] \lesssim \frac{\int_{-\varepsilon}^{\varepsilon} dh e^{-h^2}}{\int_{-\kappa}^{\kappa} dh e^{-h^2}} = \frac{\mathsf{erf}\varepsilon}{\mathsf{erf}\kappa},$$

which is again of order ε/κ for $\varepsilon\ll 1$ and $\kappa\ll 1$, (where erf denotes the standard error function).

Thus we find that adding a random correction $\kappa\boldsymbol{\xi}$ at each time instant gives typically an order ε/κ chance of landing in the region $|h|<\varepsilon$, the region where the unknown vector field $\mathbf{f}^\varepsilon$ dominates. This implies that the fraction of time $\mu(\mathbf{x})$ for which the vector field $\mathbf{f}^\varepsilon$ dominates is typically of order ε/κ. Let us therefore write $\mu(h)=\eta'(h)\varepsilon/\kappa$ where $\eta(h)$ lies between 0 and 1, then (10) becomes

$$\dot{\mathbf{x}} = \lambda(h)\mathbf{f}^+(\mathbf{x}) + (1-\lambda(h))\mathbf{f}^-(\mathbf{x}) + \frac{\varepsilon}{\kappa}\eta(h)\times \left[\mathbf{f}^\varepsilon(\mathbf{x}) - \lambda(h)\mathbf{f}^+(\mathbf{x}) - (1-\lambda(h))\mathbf{f}^-(\mathbf{x})\right] . \quad (15)$$

In the limit $\varepsilon\to 0$ we recover the simple result (3). Thus Filippov's method is justified provided that the size of the error, κ, be nonzero (or arbitrarily small provided $\varepsilon/\kappa\to 0$ as $\varepsilon\to 0$). We should also assume that the unknown function, $\mathbf{f}^\varepsilon$, which dominates near the switch, has order less than $1/\varepsilon$, which means simply that the vector field remains bounded, a reasonable assumption that requires no detailed knowledge of $\mathbf{f}^\varepsilon$ itself.

Let us illustrate all this with an everyday example from classical mechanics – an object in stick-slip contact with a surface. Attempts to form dynamical models for the basic force of dry friction between rigid bodies continues to humble modern dynamical theory. The morphology of the contact surface between moving bodies, the dynamics of microscopic irregularities and their nonlinear responses, earn this simple phenomenon a place in the modern fashion of complex systems. Most models retain some resemblance to Coulomb's "constant times normal force resisting the direction of motion": the force switches abruptly as the object changes direction.

The discontinuity in the dynamical laws of friction manifests, on the macroscopic scale at least, by destroying information. Push an object along a rough surface and you can predict if and where it will become stuck. But find an object at rest and you cannot infer whether it was previously in motion at all. Such non-invertibility is the dynamical hallmark of discontinuity.

The basic form for the friction force F felt by an object moving at speed u over a rough surface, is

$$F(u) = F_N \times \begin{cases} +\mu & \text{if } u>0\,, \\ -\mu & \text{if } u<0\,, \\ \mu_s & \text{if } u=0\,, \end{cases}$$

where μ is the coefficient of kinetic friction, μ_s is the coefficient of static friction, and F_N is the force on the object normal to the surface. Empirical studies give a value of μ_s not clearly related to, and often larger in size than, the kinetic coefficient μ. Contrast this with the Filippov method (3) which would give the coefficient of friction for $u=0$ in the narrower range $\lambda\mu+(1-\lambda)(-\mu)$ for $\lambda\in[0,1]$, i.e. in $[-\mu,+\mu]$. The fact that experiment allows $\mu_s\notin[-\mu,+\mu]$ suggests that (3) is inadequate for this case, and an added force $\mathbf{f}^\varepsilon$ as in (10) must be applied to account for the excess static friction force.

To see this friction force at work let us give our object on a surface just enough dynamics to be interesting. Let the surface move at a constant speed $v=-1$. Attach the object to a spring that exerts a force $-x$, and a damper exerting a force $-\dot{x}$. Give the object a mass $m=5$, a constant normal force $F_N=mg=50$, with a coefficient of kinetic friction $\mu\approx\frac{1}{10}$. (These values are convenient for illustration and not crucial to the results). The friction force depends on the relative speed $u=\dot{x}+1$ between the object and the surface, so we have

$$\dot{x} = u-1\,, \qquad 5\dot{u} = F(u)-x-u+1\,. \quad (16)$$

Applying the Filippov method, on $u=0$ we let $F\in[-5,5]$. The phase space trajectories can be sketched by inspection as in fig.1. On $u=0$, the force field is such

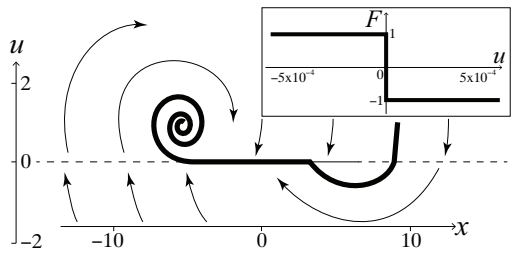

FIG. 1. Sketch of the system (16), using the Coulomb friction model shown inset, with Filippov's rule giving $F\in[-5,5]$ at $u=0$.

that solutions cross from $u>0$ to $u<0$ if $x>6$, and from $u<0$ to $u>0$ if $x<-4$. In between, where $-4<x<6$, solutions stick to the disconituity surface $u=0$ until they reach $x=-4$, whereupon they are released to spiral in towards an equilibrium at $(x,u)=(-5,1)$. In the figure we sketch a trajectory for the object starting at $x=9$ with speed $u=1$: it undergoes a change in direction as u becomes negative for a brief time, then sticks until the spring is under enough tension, at

$x = -4$, to pull the object free and attract it towards the equilibrium. Now we will simulate the same trajectory in more complex models with and without errors.

Assume that the friction law should actually be a smooth sigmoid-like function. Two toy models of friction motivated by observation (e.g.[2, 3, 8, 11, 12, 15]) are shown inset in fig.2. In both models we take $\varepsilon = 10^{-4}$ (the precise functions taken have little bearing on the numerical results). Simulations for these two friction models are shown in fig.2, depicting a trajectory over time $t = 26$ with initial conditions as in fig.1.

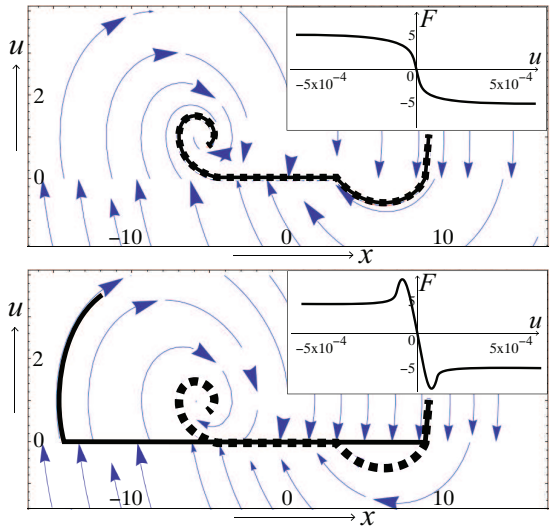

FIG. 2. Simulation with error (dotted path) and without error (full curve), using friction models $F(u) = -\tanh(u/\varepsilon)$ (top) and $F(u) = -(1 + 2\pi e^{-u^2/\varepsilon^2}) \tanh(u/\varepsilon)$ (bottom) shown inset.

The full curves show solutions of (16) for the two different smoothed friction models. The first model closely agrees with the Filippov method in fig.1. In the second model the object becomes stuck to the surface earlier and for a longer time, requiring a greater tension in the spring to pull it free at $x \approx -14$. The object can stick more than once before going to the equilibrium, contrary to the first model where only one sticking phase is possible.

The dotted curves show the same solution, except that we add a random vector $(\kappa\xi_x, \kappa\xi_u)$ to (x, u), applied after every interval $\Delta t = 0.26$, with $\xi_x^2 + \xi_u^2 = 1$ and $\kappa = 5 \times 10^{-3}$. For the first model in fig.2 this error has little effect. In the second model the error removes the extended sticking phase, restoring the Filippov-type dynamics seen in fig.1. So in both cases, with sufficient error $\kappa \gtrsim \varepsilon$ the simple system (3) is valid near $u = 0$.

Another practical example is an electronic control circuit. The dynamics of ideal 'sliding-modes' is found by applying Utkin's method of equivalent control [14] to (1), giving (5). When $\mathbf{f}$ depends nonlinearly on c this is not equivalent to Filippov's result (3). But the results here show that unmodelled errors of size κ will tend to push the observed dynamics towards Filippov's dynamics *if* the switch takes place over a region of size $\varepsilon < \kappa$. The pioneers of nonsmooth dynamics suggested such a result without formulating it in any detail (see e.g. [1]). It is hoped that the present letter allows this argument to be developed formally, with more rigorous treatments considering how distributions of noise, hysteresis, delay, and other perturbations evolve near the discontinuity.

In the example above we smoothed out the friction law to facilitate simulation. The computational and analytic effects of smoothing are actually poorly understood. Indeed, we have shown that any system (1) may be the limit of a system (10), the latter representing infinitely many possibilities because $\mathbf{f}^\varepsilon$ is unknown. The fate of (3) under singular perturbation has been of recent interest, see e.g. [13], but no attention has been given to the effect of μ, λ, or $\mathbf{f}^\varepsilon$, in the general regularization (10).

The idea presented here highlights a danger of overmodeling complex behaviour. Unknown errors can cause a system to behave more like a crude model (3) than a more refined one (10). Stated another way, discontinuous models owe their unreasonable effectiveness to unmodelled errors that cancel themselves out. By analyzing the ambiguity in how we treat the discontinuity we can quantify the possible effect of such errors, and determine when they can be neglected.

[1] A. A. Andronov, A. A. Vitt, and S. E. Khaikin. *Theory of oscillations*. Moscos: Fizmatgiz (in Russian), 1959.

[2] J. Awrejcewicz, L. Dzyubak, and C. Grebogi. *Nonlinear Dynamics*, 42:383–394, 2005.

[3] P. R. Dahl. A solid friction model. *TOR-158(3107-18), The Aerospace Corporation, El Segundo, CA*, 1968.

[4] M. di Bernardo, C. J. Budd, A. R. Champneys, and P. Kowalczyk. *Piecewise-Smooth Dynamical Systems: Theory and Applications*. Springer, 2008.

[5] A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*. Kluwer Academic Publ., 1988.

[6] M. R. Jeffrey. *PRL*, 106(254103):1–4, 2011.

[7] O. Makarenkov and J. S. W. et al Lamb. *Physica D Special Issue*, 241(22):1825–2082, 2012.

[8] A.J. McMillan. *Journal of Sound and Vibration*, 205(3):323 – 335, 1997.

[9] D. J. W. Simpson and R. Kuske. *arxiv.org/abs/1204.5792*, 2012.

[10] J-J. E. Slotine and W. Li. *Applied Nonlinear Control*. Prentice Hall, 1991.

[11] R. W. Stark, G. Schitter, and S. Stemmer. *Ultramicroscopy*, 100(3–4):309 – 317, 2004.

[12] D. E. Stewart. *SIAM Review*, 42(1):3–39, 2000.

[13] M. A. Teixeira, J. Llibre, and P. R. da Silva. *J.Dyn.Diff.Equat.*, 19(2):309–331, 2007.

[14] V. I. Utkin. *IEEE Trans. Automat. Contr.*, 22, 1977.

[15] J. Wojewoda, S. Andrzej, M. Wiercigroch, and T. Kapitaniak. *Phil. Trans. R. Soc. A*, 366:747–765, 2008.