Physica D 269 (2014) 76–85

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Smoothing non-smooth systems with low-pass filters

James Hook *

Alan Turing Building, Oxford Rd, Manchester, UK

HIGHLIGHTS

- We introduce the notion of smoothing a non-smooth system.
- Our smoothing is equivalent to filtering time-series data.
- The smoothing is useful for computing stability.

ARTICLE INFO

Article history:
Received 23 July 2013
Received in revised form
31 October 2013
Accepted 23 November 2013
Available online 3 December 2013
Communicated by G. Stepan

Keywords:
Non-smooth dynamics
Regularisation
Time-series
Filtering
Lyapunov stability

ABSTRACT

Low pass filters, which are used to remove high frequency noise from time series data, smooth the signals they are applied to. In this paper we examine the action of low pass filters on discontinuous or non-differentiable signals from non-smooth dynamical systems. We show that the application of such a filter is equivalent to a change of variables, which transforms the non-smooth system into a smooth one. We examine this smoothing action on a variety of examples and demonstrate how it is useful in the calculation of a non-smooth system's Lyapunov spectrum.

0. Introduction

Non-smooth dynamical systems are used to model mechanical systems with impacts or friction, as well as control systems with switching between distinct modes of operation. Non-smooth systems are also interesting mathematically as they generically exhibit bifurcation structures that would be impossible or of high co-dimension in the space of smooth systems [1].

In this paper we introduce the notion of smoothing a non-smooth system with a low-pass filter. The idea is that the filter's action on the time-series can be used to construct a change of variables that transforms a non-smooth system into a smooth one. There are some subtleties here, the 'smoothed' system will not be smooth everywhere as singular discontinuities (grazes and chattering points) will be mapped to singularities in the new flow, also the transformation and the smoothed system will typically be impossible to compute analytically. However we will still be able to calculate them for simple examples or approximate them numerically for more complicated systems. To apply a smoothing transformation numerically to an orbit we simply apply the associated low pass filter to the time-series. Indeed whenever an engineer analyses data from a non-smooth system that has been filtered they are inadvertently studying a 'smoothed' system of the sort presented here.

There are therefore two complimentary reasons for trying to understand the action of these smoothing transformations. Firstly we might find the smoothing action useful or interesting in its own right (we will show that it is useful for computing Lyapunov exponents) and secondly such systems are already being investigated whenever experimental data is smoothed with a low pass filter.

This paper is organised as follows. In Section 1 we show how a non-smooth system can be transformed into a smooth one using a change of variables. The approach used in Section 1 is rather ad hoc and does not use low pass filters but it allows us to understand the link between the smooth and the non-smooth system in as simple a setting as possible. In Section 2 we show how a similar smoothing action can be achieved using our low-pass filter formulation. We examine some analytic examples and consider some of the signal processing issues associated with the transformation. In Section 3 we briefly explain the state space reconstruction method which enables us to model a differentiable dynamical system from its time-series data and examine some simple numerical examples. In Section 4 we argue that linear stability (when it exists) is preserved by the smoothing procedure. In Section 5 we apply

* Tel.: +44 7817975372.
E-mail address: james.hook@manchester.ac.uk.

0167-2789/$ – see front matter © 2014 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.physd.2013.11.016

the smoothing procedure to time series from an impacting Duffing oscillator and calculate its Lyapunov spectrum using a time-series method that relies on differentiability.

1. Ad hoc smoothing

Consider a mass on a linear spring whose motion is obstructed by a wall placed at the spring's natural length. Suppose that when the mass hits the wall it bounces off it elastically with coefficient of restitution c. Let $x(t)$ measure the distance from the wall to the mass. The motion of the mass is governed by $\ddot{x} = -x$, along with the rule that whenever $\lim_{\tau\to t} x(\tau) = 0$ we set $\dot{x}(t) = \lim_{\tau\to t} -c\dot{x}(\tau)$. The state space of this system is therefore $X = \mathbb{R}_+ \times \mathbb{R}$.

Orbits to this system comprise of a series of smaller and smaller semi-circles, see Fig. 1. A solution evolves by describing one of the semi-circles until it reaches $x = 0$, when it instantaneously jumps to the start of a new smaller semi-circle and so on. This roughly periodic behaviour is just like that of a smooth system with a stable fixed point, where solutions spiral into the equilibrium. Indeed, we can imagine sticking a pin into the origin of this picture and stretching the space around it to fill the plane. The two sides of the boundary would meet and it would be possible to glue them together so that the jump 'take offs' and 'landings' joined together. If any kinks in the picture could then be ironed out we would have something that looked exactly like the stable equilibrium of a smooth system.

It turns out that for this simple example we can formulate a transformation that has these exact properties.

$$T\left(\begin{bmatrix} r\sin(\theta) \\ r\cos(\theta)\end{bmatrix}\right) = \begin{bmatrix} rc^{\frac{\theta}{\pi}}\sin(2\theta) \\ rc^{\frac{\theta}{\pi}}\cos(2\theta)\end{bmatrix},$$

maps the semi-circle starting at $[0, \dot{x}]^t$ to the 360° spiral starting at $[0, \dot{x}]^t$ and finishing at $[0, c\dot{x}]^t$. T transforms the original non-smooth system to a smooth system governed by

$$\begin{bmatrix}\dot{p} \\ \dot{q}\end{bmatrix} = \begin{bmatrix} \dfrac{\log(c)}{\pi} & 1 \\ -1 & \dfrac{\log(c)}{\pi}\end{bmatrix}\begin{bmatrix} p \\ q\end{bmatrix}. \tag{1}$$

Let $\phi : X \times \mathbb{R}_+ \mapsto X$ be the flow of the original discontinuous system and $\varphi : \mathbb{R}^2 \times \mathbb{R}_+ \mapsto \mathbb{R}^2$ be the flow of the new smoothed system. The transformation T provides the commutation

$$\phi_t([x, \dot{x}]^t) = T^{-1} \circ \varphi_t \circ T([x, \dot{x}]^t),$$

so that we can substitute one flow for another. Likewise their stability is related by

$$\left.\frac{d\phi_t([x', \dot{x}']^t)}{d[x', \dot{x}']^t}\right|_{[x,\dot{x}]^t} = \left.\frac{dT^{-1}([p, q]^t)}{d[p, q]^t}\right|_{\varphi_t\circ T([x,\dot{x}]^t)} \times \left.\frac{d\varphi_t([p, q]^t)}{d[p, q]^t}\right|_{T([x,\dot{x}]^t)} \times \left.\frac{dT([x', \dot{x}']^t)}{d[x', \dot{x}']^t}\right|_{[x,\dot{x}]^t}.$$

This alternative expression is much simpler to evaluate as we no longer have to worry about repeated application of saltation matrices every time the orbit crosses the discontinuity. Instead we only need to evaluate the stability of the smooth flow and multiply it by the derivatives of T. Moreover since the derivatives of T and its inverse are everywhere bounded the Lyapunov spectrums of the two systems are identical. It is easy to show from the smoothed system that both Lyapunov exponents are equal to $\log(c)/\pi$.

We are able to play the same game with the bouncing ball system. This evolves according to $\ddot{x} = -g$ along with the rule that whenever $\lim_{\tau\to t} x(\tau) = 0$, we set $\dot{x}(t) = \lim_{\tau\to t} -c\dot{x}(\tau)$. Here the analysis is only a little more complicated, we start by mapping parabolas to 360° spirals with a transformation given by

$$T\left(\begin{bmatrix} x \\ y\end{bmatrix}\right) = \begin{bmatrix} \sqrt{2E} - \dfrac{(1-r)(\sqrt{2E}-x)}{2} \\ \dfrac{2\pi}{\log(c)}\log\left(1 - \dfrac{(1-c)(\sqrt{2E}-x)}{2\sqrt{2E}}\right)\end{bmatrix},$$

where $E = \frac{x^2}{2} + gy$ is the energy. This transformation provides a conjugacy between the bouncing ball system and the smooth (everywhere except the origin) system

$$\begin{bmatrix}\dot{p} \\ \dot{q}\end{bmatrix} = \frac{2\pi g(c-1)}{\log(c)\sqrt{p^2+q^2}}\begin{bmatrix} \dfrac{\log(c)}{2\pi} & 1 \\ -1 & \dfrac{\log(c)}{2\pi}\end{bmatrix}\begin{bmatrix} p \\ q\end{bmatrix}. \tag{2}$$

It should come as no surprise that this 'smoothed' system is not quite smooth. The conjugacy between its flow and that of the bouncing ball means they must share stability properties and the bouncing ball is a singular system; all orbits reach the origin in finite time.

It would be fantastic if we could explicitly construct such conjugacies for more complicated non-smooth systems. Unfortunately although our filter based transformation has the desired smoothing action, explicitly applying it to obtain the smoothed system is not typically possible as it requires integrating orbits to the original system. However we will show in the next section that it is still possible to examine the smoothed system by smoothing time-series data recorded from the non-smooth system.

2. Smoothing with low-pass filters

A finite impulse response filter Ψ is a linear operator given by

$$\Psi(f)(t) = \int_{-w}^{0} f(t+\tau)h(\tau)d\tau, \tag{3}$$

where $h(\tau)$ is the kernel and w the window. We can use filters like these to create a smoothing transformation. Let $\phi : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}^n$ be the flow of a non-smooth system with state space $\mathbb{R}^n$. Now define $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ by

$$[T(\underline{x})]_i = \int_{-w}^{0} [\phi_\tau(\underline{x})]_i h(\tau)d\tau, \tag{4}$$

so that $T(\underline{x})$'s ith component is calculated by integrating the value of the ith co-ordinate of $\underline{x}$'s orbit up to w seconds backwards in time. For the time-being we will assume that T is an invertible map. Given this assumption T induces a new flow φ defined by

$$\varphi_t(\underline{p}) = T \circ \phi_t \circ T^{-1}(\underline{p}). \tag{5}$$

So that if we think of $[\phi_t(\underline{x})]_i$ as a function of time t and likewise for φ, then we have

$$\begin{tikzcd} \underline{x} \arrow[r, "T"] \arrow[d, "\phi"'] & T(\underline{x}) \arrow[d, "\varphi"] \\ \phi(\underline{x}) \arrow[r, "\Psi"'] & \Psi[\phi(\underline{x})] = \varphi[T(\underline{x})] \end{tikzcd} \tag{6}$$

Or in words, the unfiltered orbits from the transformed (smoothed) system are identical to the filtered orbits of the original system. Whenever we analyse time-series data from a non-smooth system that has been filtered (to reduce noise say) we are inadvertently studying one of these smoothed systems. It is therefore important to understand how this smoothing process affects the features that we are interested in e.g. stability and grazing points.

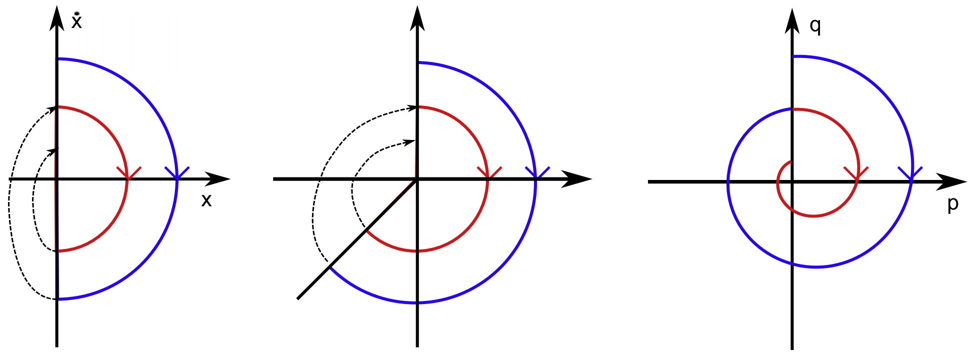

Fig. 1. Left to right: deforming the discontinuous system into a smooth system.

2.1. Invertibility

In order to be able to reconstruct a smoothed system it is essential that T be invertible. This requirement can be split into two parts. Firstly we require the filter Ψ to be invertible as an operator on the time series $\phi_t(x)$. Secondly given this first condition we still require that the transformation T itself is an injection. Both of these problems are well studied in the context of smooth systems [2,3].

Filters, as defined in (3), are best described in Fourier space where we have

$$\widehat{\Psi(f)}(s) = \widehat{f}(s) \times H(s),$$

where $\widehat{\ }$ denotes the Fourier transform and H is the transfer function

$$H(s) = \int_0^w h(-\tau)\exp[-2\pi i s\tau]d\tau.$$

If the application of a filter is to be invertible it is essential that $H(s) = 0$ only when $\widehat{f}(s) = 0$ also. Of course if we are applying a filter to remove noise the transformation will not be invertible (we cannot expect to recover deleted noise) but we must still compare the spectrum of the input data to the transfer function of the filter to ensure that the only information lost is in a band consigned as noise.

As we will show in Example 4 invertibility of Ψ does not guarantee invertibility of T. In these cases we use the method of delays to construct an invertible map with the required properties. Takens' theorem states that for generic smooth flow ϕ, delay d and smooth measuring function f, there exists finite m such that

$$F(\underline{x}) = \left[f(\underline{x}), f[\phi_{-d}(\underline{x})], \ldots, f[\phi_{-md}(\underline{x})]\right],$$

is invertible and provides the conjugacy between ϕ and a diffeomorphically equivalent system φ. Of course our systems are non-smooth and our measuring function depend on the flow so they are not generic. Therefore Takens' theorem gives us no guarantees but we follow its spirit and find that the method of delays works well in this non-smooth setting.

Feeny [4] investigates the related problem of reconstructing a non-smooth system's state space from its time series. Feeny shows that this is not always possible for systems with a sliding mode. However this scenario only occurs when a strict subset of the state variables are used, whereas here we are talking about using every single variable to produce the smoothed system.

2.2. Differentiability

Recall that the transformed system's time-evolution map is

$$\varphi_t(\underline{p}) = T \circ \phi_t \circ T^{-1}(\underline{p}),$$

where ϕ is the time-evolution map of the original system. Let $\underline{x}^*$ be a point in the state space of the original system, which is contained in an open set containing no discontinuities, so that for $\underline{x}$ close to $\underline{x}^*$ we have

$$\frac{d\phi_t(\underline{x})}{dt} = F(\underline{x}),$$

for some smooth vector field F. Then the smoothed system is the ODE

$$\frac{d\varphi_t(\underline{p})}{dt} = \frac{dT}{d\underline{x}}F[T^{-1}(\underline{p})], \tag{7}$$

in the vicinity of $T(\underline{x})$.

Therefore, in order for the transformed system to be differentiable it is necessary that T be invertible, and a sufficient condition is that T be differentiable and the flow in the original system be differentiable also. In the subsequent examples we will show that this condition can be relaxed to allow non-grazing jumps, switches and slides.

The smoothing transformation T will be differentiable at a point $\underline{x}$, if $\underline{x}$'s w second backwards time orbit is not in the neighbourhood of any grazes, cusps or slides—although we will show in Example 3 that it is possible to smooth a non-grazing slide discontinuity.

Example 1 (*Moving Average Transformation Applied Twice to a Discontinuous System*). Let ϕ be the flow of the system governed by $[\dot{x}, \dot{y}]^t = [1, 0]^t$ along with the rule that whenever $\lim_{\tau\to t} x(\tau) = 0$, we set $[x(t), y(t)]^t = [1, \lim_{\tau\to t} 2y]^t$, see Fig. 2. In this and the subsequent analytic examples we will use the moving average transformation T, which is the simplest of our smoothing transformations with unit window $w = 1$ and constant kernel $h = 1$. This choice is purely to allow some analytic tractability.

The moving average transformation is piecewise smooth on the state space of our system. For $x < 0$ or $x \geq 2$ there are no discontinuities in the one-second backward time flow so that

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \int_{-1}^{0} \begin{bmatrix} x+\tau \\ y \end{bmatrix} d\tau = \begin{bmatrix} x - \dfrac{1}{2} \\ y \end{bmatrix}.$$

For $1 \leq x < 2$ there will be a discontinuity in the one-second backwards time flow so that the integral expression for T will contain contributions from before and after the jump

$$\begin{aligned} T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) &= \int_{-1}^{1-x} \begin{bmatrix} x-1+\tau \\ \dfrac{y}{2} \end{bmatrix} d\tau + \int_{1-x}^{0} \begin{bmatrix} x+\tau \\ y \end{bmatrix} d\tau \\ &= \begin{bmatrix} 2x - \dfrac{5}{2} \\ \dfrac{xy}{2} \end{bmatrix}. \end{aligned} \tag{8}$$

The transformed system φ is governed by the non-smooth ODE

$$\begin{bmatrix} \dot{p} \\ \dot{q} \end{bmatrix} = \frac{dT}{d[x, y]^t} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}_{T^{-1}[p,q]^t}, \tag{9}$$

inverting (8) and substituting into (9) we obtain

$$\begin{bmatrix}\dot p\\ \dot q\end{bmatrix} = \begin{cases}\begin{bmatrix}1\\0\end{bmatrix}, & \text{for } p < -0.5,\\ \begin{bmatrix}2\\ \dfrac{2q}{p+\frac{5}{2}}\end{bmatrix}, & \text{for } -0.5 \le p < 1.5,\\ \begin{bmatrix}1\\0\end{bmatrix}, & \text{for } 1.5 \le p.\end{cases} \tag{10}$$

Applying the moving average transformation we have obtained a non-differentiable but continuous system conjugate to our original discontinuous system.

We can now apply the transformation a second time, which is equivalent to one application of a smoother filter with window $w = 2$ and kernel

$$h(\tau) = \begin{cases}\tau + 2, & \text{for } -2 \le \tau < -1,\\ -\tau, & \text{for } -1 \le \tau \le 0.\end{cases}$$

Again, this transformation is piecewise smooth but now depends on the behaviour of the two second backward time flow,

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{cases}\begin{bmatrix}x-1\\y\end{bmatrix}, & \text{for } x < 0,\\ \begin{bmatrix}\dfrac{x^2}{2} - \dfrac{3}{2}\\ \left(\dfrac{(x-1)^2}{4} + \dfrac{1}{2}\right)y\end{bmatrix}, & \text{for } 1 \le x < 2,\\ \begin{bmatrix}-\dfrac{x^2}{2} + 4x - \dfrac{11}{2}\\ \left(1 - \dfrac{(x-3)^2}{4}\right)y\end{bmatrix}, & \text{for } 2 \le x < 3,\\ \begin{bmatrix}x-1\\y\end{bmatrix}, & \text{for } 3 \le x.\end{cases}$$

As with the previous transformation we explicitly invert the transformation T then substitute it (9). The discontinuous system is transformed into the differentiable system governed by

$$\begin{bmatrix}\dot a\\ \dot b\end{bmatrix} = \begin{cases}\begin{bmatrix}1\\0\end{bmatrix}, & \text{for } a < -1,\\ \begin{bmatrix}\sqrt{2a+3}\\ \dfrac{4b(\sqrt{2a+3}-1)}{(\sqrt{2a+3}-1)^2+2}\end{bmatrix}, & \text{for } -1 \le a < 0.5,\\ \begin{bmatrix}\sqrt{5-2a}\\ \dfrac{2b(\sqrt{5-2a}-1)}{4-(1-\sqrt{5-2a})^2}\end{bmatrix}, & \text{for } 0.5 \le a < 2,\\ \begin{bmatrix}1\\0\end{bmatrix}, & \text{for } 2 \le a.\end{cases} \tag{11}$$

Through double application of the moving average transformation we have obtained a differentiable system conjugate to our original discontinuous system. The non-differentiable features that would have affected the stability of the system have been smoothed out. But their effect on the dynamics has not been lost, they have been integrated into the new smooth flow. Smoothing non-smooth systems does not destroy information about the discontinuities, rather it encodes this information in a different way. This smoothing action at regular discontinuities (jumps or switches that are not grazes or chattering points) can be shown to work in a general setting. In [5] we present normal forms for smoothing these sorts of discontinuities with the moving average transformation.

Example 2 (*Moving Average Transformation Applied Once to a Non-Differentiable Graze*)**.** Let ϕ be the flow governed by

$$\begin{bmatrix}\dot x\\ \dot y\end{bmatrix} = \begin{cases}\begin{bmatrix}0\\-1\end{bmatrix}, & \text{for } x < y^2,\\ \begin{bmatrix}1\\0\end{bmatrix}, & \text{for } x \ge y^2.\end{cases}$$

This system has a grazing point at the origin, see Fig. 3. The moving average transformation is given by

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{cases}\begin{bmatrix}x\\ y-\dfrac{1}{2}\end{bmatrix}, & \text{for } x < y^2,\\ \begin{bmatrix}y^2 + \dfrac{(x-y^2)^2}{2}\\ y + \dfrac{(1-x+y^2)^2}{2}\end{bmatrix}, & \text{for } y^2 \le x < y^2+1,\\ \begin{bmatrix}x - \dfrac{1}{2}\\ y\end{bmatrix}, & \text{for } y^2 + 1 \le x.\end{cases}$$

Unfortunately we cannot explicitly invert T as it requires us to solve a quartic polynomial. However we can approximate the flow at the image of the graze by ignoring terms that are smaller than x or y^2 to obtain a local description of the transformed system for $[p, q]^t \approx [0, 0.5]^t$ given by

$$\begin{bmatrix}\dot p\\ \dot q\end{bmatrix} = \begin{bmatrix}\sqrt{p}\\ \sqrt{p}\end{bmatrix} + \begin{bmatrix}0\\-1\end{bmatrix} + \begin{bmatrix}0 & -1\\ 0 & -1\end{bmatrix}\begin{bmatrix}p\\ q-\dfrac{1}{2}\end{bmatrix}. \tag{12}$$

Notice that although the square root expression will be non-differentiable at $p = 0$ the affine part of the right hand side ensures that there is no problem with non-uniqueness of solution.

It is easy to show that the transformed system is differentiable everywhere except at the image of the grazing orbit. Away from this orbit the smoothing action will be exactly as in Example 1, the non-differentiable switch will be replaced with a differentiable switch and any saltation associated with the switch integrated into the new flow. As with the regular jump and switch this smoothing action is totally general and we derive a normal form for the moving average smoothed graze in [5].

Example 3 (*Moving Average Transformation Applied Once to a Non-Differentiable Slide*)**.** Let ϕ be the flow governed by

$$\begin{bmatrix}\dot x\\ \dot y\end{bmatrix} = \begin{cases}\begin{bmatrix}1\\-1\end{bmatrix}, & \text{for } y > 0,\\ \begin{bmatrix}1\\1\end{bmatrix}, & \text{for } y < 0.\end{cases}$$

This system has a slide discontinuity on the set $\{[x, y]^t : y = 0\}$. To extend the vector field to this set, which we call the sliding surface, we use the Filippov differential inclusion method [6], to obtain $[\dot x, \dot y]^t = H(x) = [1, 0]^t$ on $y = 0$.

Orbits to this system switch direction non-differentiably as they join the sliding surface $y = 0$, then remain on this set for the rest of time, see Fig. 4. As a result the moving average transformation is not uniquely defined for points on the sliding surface, as they have many possible histories to average over. Our approach is to

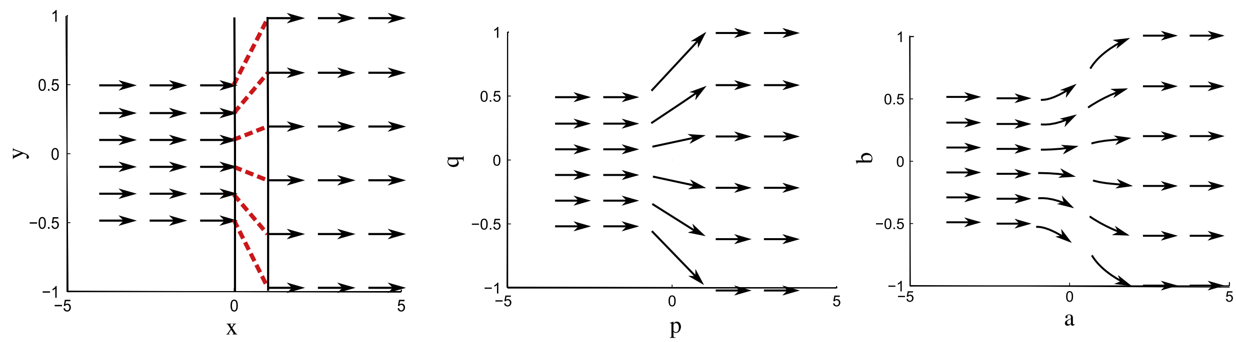

Fig. 2. Left to right: vector field of original discontinuous system, once smoothed system and twice smoothed system.

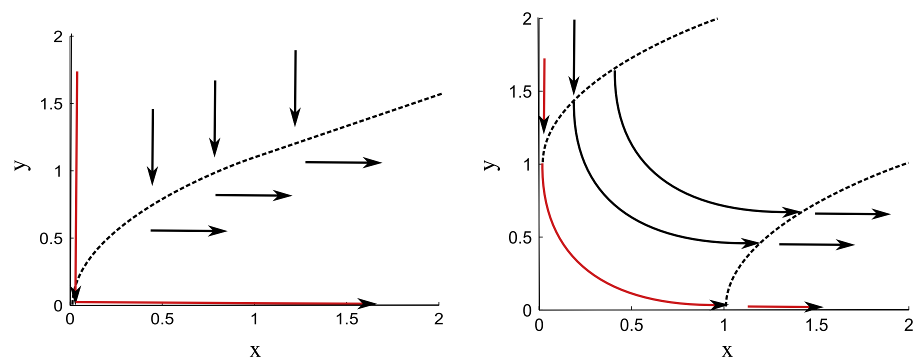

Fig. 3. Left to right: vector field of original discontinuous system and once smoothed system. The red curves (closest to axis) indicated the grazing orbit and its T image.

take every possible branch of the transformation. Away from the sliding surface the moving average transformation is given by

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{cases}\begin{bmatrix}x-\dfrac{1}{2}\\y+\dfrac{1}{2}\end{bmatrix}, & \text{for } y>0,\\[2ex] \begin{bmatrix}x-\dfrac{1}{2}\\y-\dfrac{1}{2}\end{bmatrix}, & \text{for } y<0.\end{cases}$$

For a point $(x, 0)^t$ on the sliding surface

$$T\left(\begin{bmatrix}x\\0\end{bmatrix}\right)=\int_{-1}^{0}\phi_\tau\left(\begin{bmatrix}x\\0\end{bmatrix}\right)dx,$$

has two continuous families of branches. For all $a \in [x, \infty)$ it is possible that x's backwards orbit leaves the sliding surface in the positive or negative y direction at $(a, 0)$. We define the branches of T by

$$T_{a,\pm}\left(\begin{bmatrix}x\\0\end{bmatrix}\right)=\begin{cases}\begin{bmatrix}x-\dfrac{1}{2}\\0\end{bmatrix}, & \text{for } 1<x-a,\\[2ex] \begin{bmatrix}x-\dfrac{1}{2}\\ \dfrac{\pm(x-a)^2}{2}\end{bmatrix}, & \text{for } 0\le x-a\le 1.\end{cases}$$

From (9), each branch contributes to the smoothed system by

$$\begin{bmatrix}\dot p\\\dot q\end{bmatrix}=\frac{dT_{a,\pm}}{dx}H_x\left(T_{a,\pm}^{-1}\left(\begin{bmatrix}p\\q\end{bmatrix}\right)\right),\quad \text{on } \begin{bmatrix}p\\q\end{bmatrix}=T_{a,\pm}\left(\begin{bmatrix}x\\0\end{bmatrix}\right).$$

From $1 < x - a$ we have $[\dot p, \dot q]^t = [1, 0]^t$ on $q = 0$, and from $0 \le x - a \le 1$ we have

$$\begin{bmatrix}\dot p\\\dot q\end{bmatrix}=\begin{bmatrix}1\\\pm(x-a)\end{bmatrix}\times 1,$$

on $q = \pm(x - a)^2/2$. So that the smoothed system is given by the ODE

$$\begin{bmatrix}\dot p\\\dot q\end{bmatrix}=\begin{cases}\begin{bmatrix}1\\-1\end{bmatrix}, & \text{for } q>0.5,\\ \begin{bmatrix}1\\-\sqrt{2q}\end{bmatrix}, & \text{for } 0\le q\le 0.5,\\ \begin{bmatrix}1\\\sqrt{-2q}\end{bmatrix}, & \text{for } -0.5\le q\le 0.\\ \begin{bmatrix}1\\1\end{bmatrix}, & \text{for } q<-0.5.\end{cases} \tag{13}$$

By taking all possible branches of T we have expanded the sliding surface into a strip of orbits which merge into $q = 0$ in a differentiable manner through a stable square root term in the flow, see Fig. 4. Orbits which have joined the sliding surface within the last second are differentiated from each other by their unique histories, but after one second their one second histories are all the same and their T images join the smoothed sliding surface $q = 0$. As before this smoothing action is totally general and we derive a normal form for the moving average smoothed slide in [5].

Example 4 (*Bump Transformation Applied to a Noisy Data From a Discontinuous System*). We return to the mass on a spring system introduced in Section 1. The motion of the mass is governed by $\ddot x = -x$ along with the rule that whenever $\lim_{\tau\to t} x(\tau) = 0$, we set $\dot x(t) = \lim_{\tau\to t} -c\dot x(\tau)$. To set up an artificial study of noisy experimental data we simulate this system on a computer then add Gaussian noise to every variable to produce a noisy time series. The

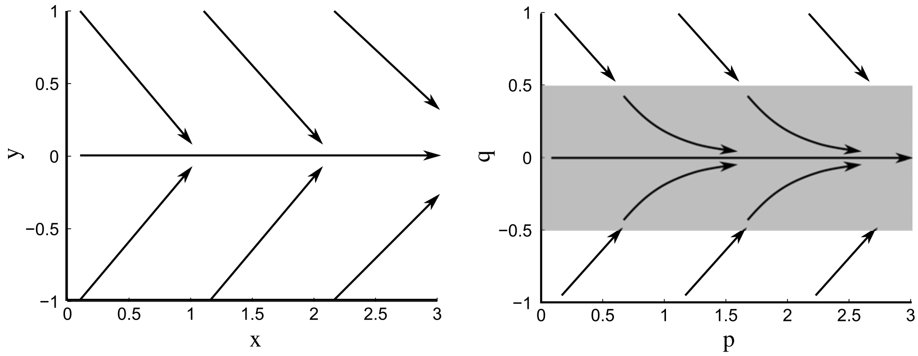

Fig. 4. Left to right: vector field of original discontinuous system and once smoothed system. The grey region represents the Φ image of the sliding surface.

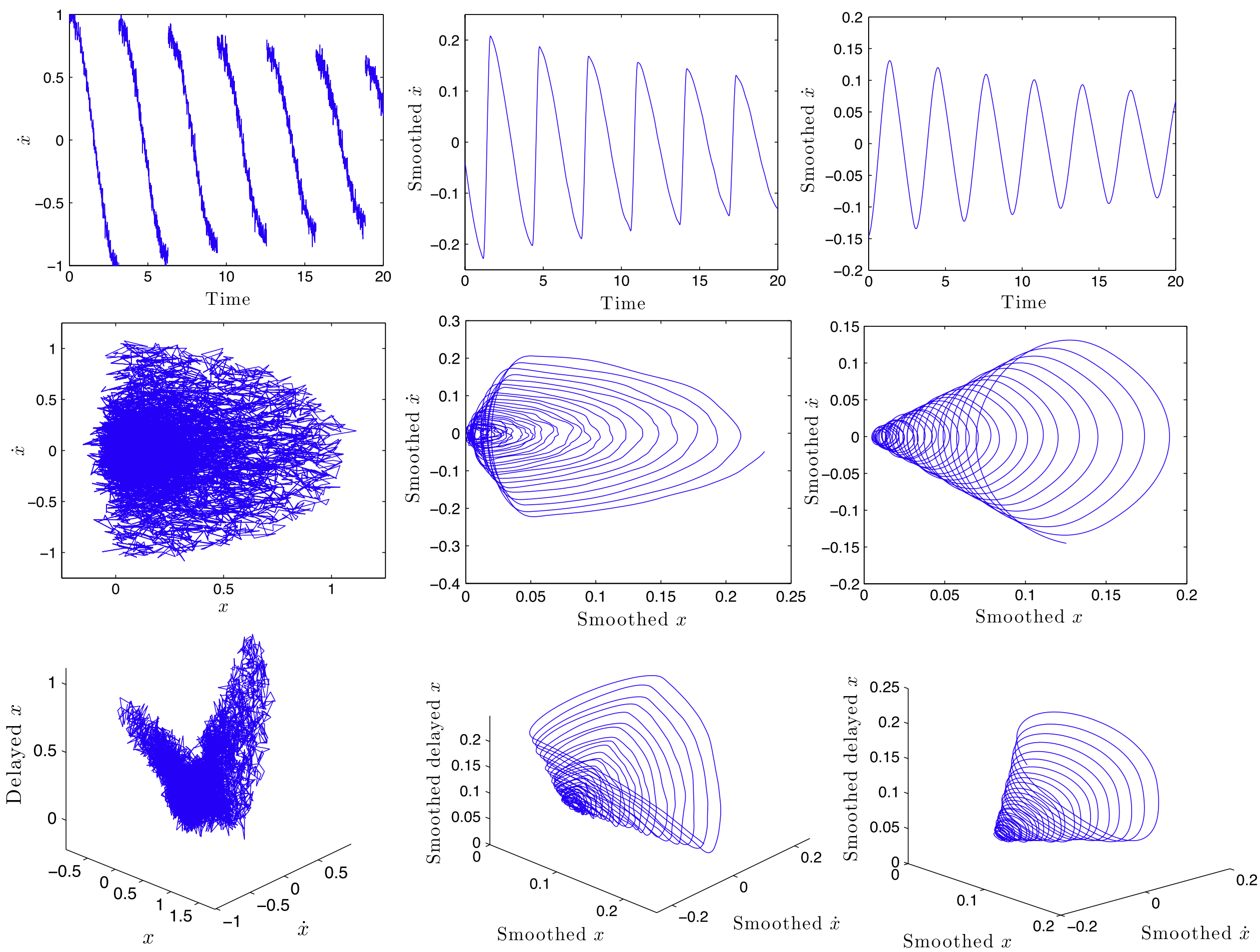

Fig. 5. Left to right: non-smooth system, bump smoothed system with $w = 0.5$ s, bump smoothed system with $w = 2$ s. Top to bottom: velocity variable versus time, phase space, delay space.

noisy data is useless for reconstructing the phase space of the system, all we see is a noisy blob, see the left column of Fig. 5.

In order to reduce the noise in the signal and smooth the discontinuity we will apply a low pass filter to the time-series. As we saw in Example 1, using a filter with a smoother kernel results in a smoother system. To obtain a totally smooth system (or at least a system with smooth orbits) we apply a filter with smooth kernel. The bump filter has smooth kernel

$$h(\tau) = e^{\frac{-1}{\tau(w-\tau)}}. \tag{14}$$

In principle this filter can completely smooth any integrable data. However, we are working in a numerical setting where we have the time series stored as a discrete sequence, so that any filter will be in the form of a weighted sum rather than an integral. Applying the discrete bump filter we expect to see any jumps in the first few derivatives to be removed. The exact action will depend on the stiffness of the original data the time step and the length of the filtering window w.

For our experiment we use two window lengths for the filter $w = 0.5$ s and $w = 2$ s. Comparing the time series of the original and smoothed systems we see that the discontinuous noisy data is transformed to smooth less noisy data, see top row of Fig. 5. The time series filtered with the shorter window is still quite stiff, this is because the discontinuity is only 'spread out' over a short window of time. The time series filtered with the shorter window is also a little noisy compared to that filtered with the longer window $w = 2$ s.

If we try to use this smoothed data to reconstruct the original phase space we have two problems. Firstly, as we would expect the data is no-longer discontinuous. Smoothing the data is equivalent to smoothing the system and we have already shown that smoothing will map a discontinuous system to a continuous one. Secondly, the transformation is not an injection; different points in the original phase space are mapped to the same point under the smoothing, so we cannot reconstruct a smooth dynamical system from this data, see the middle row of Fig. 5. To remedy this problem we add a delay vector

$$D[x(t), \dot{x}(t)] = [x(t), \dot{x}(t), x(t-d)].$$

The delayed data gives a reconstruction of the state space of the smooth system conjugate to the original system via the bump smoothing transformation with delay, see the bottom row of Fig. 5. The smoothed systems obtained with the different filtering windows are topologically equivalent to each other, but not to the original discontinuous system. The system obtained using the longer window $w = 2$ s is less stiff and its data is less noisy. We show how to choose the optimal filtering window for state space reconstructions in the next section. Note that if we were to position ourselves to 'look down' into the $w = 2$ s cone (bottom right) we would see a picture exactly like the smoothed system described in Section 1.

We therefore find that it is impossible to apply a filter to reduce noise from a non-smooth time series without transforming the system into a smooth one. However this need not be a problem. We have already shown how non-smooth discontinuities are smoothed out by the transformations, so that if we are interested in looking for e.g. a graze we know to look for a square root singularity in the flow. We argue in Section 4 that the smoothed system will have the same stability properties as the original system.

3. State space reconstructions

We have shown that the filtered data from a non-smooth system resembles unfiltered data from its smoothed system. Therefore in cases where we are unable to study the smoothed systems directly by explicitly computing and applying the transformation, we can instead record time-series data from the non-smooth system and filter it to obtain a time-series that is equivalent to a recording from the smoothed system. In this section we give a very brief and informal account of the method of state space reconstruction through which one is able to construct a numerical model of a differentiable system from its time-series. For a more thorough exposition see [7].

Suppose that ϕ is a differentiable flow on a manifold $M \subset \mathbb{R}^n$. We are able to record an orbit of the system on a computer by storing the value of $\phi_t(\underline{x})$ every τ seconds. The data is a long sequence $[\underline{x}(i)]_{i=1}^N$ with the property that $\underline{x}(i+1) = \phi_\tau[\underline{x}(i)]$.

Provided the stored orbit explores the manifold M sufficiently thoroughly we will be able to build a piecewise affine model of M and the time-τ map $\phi_\tau : M \mapsto M$ (see Fig. 6).

To construct an affine model of M and ϕ_τ in the vicinity of the point $\underline{x}(i)$ we start by searching the time-series for other data points that are within a distance ϵ of $\underline{x}(i)$. Let $\underline{x}(j_1), \underline{x}(j_2), \ldots, \underline{x}(j_m)$ be the points that are within a distance ϵ of $\underline{x}(i)$. A basis for the tangent space of M at $\underline{x}(i)$ is approximated by the span of the matrix

$$N(i) = \left[\underline{x}(j_1) - \underline{x}(i)\, \underline{x}(j_2) - \underline{x}(i) \cdots \underline{x}(j_m) - \underline{x}(i)\right],$$

which is computed by taking the SVD decomposition and ignoring basis vectors corresponding to small singular values,

$$N(i) = U(i)S(i)V(i),$$

so the basis of M's tangent space at $\underline{x}(i)$ is given by $[U_1(i), U_2(i), \ldots, U_d(i)]$ say, where $U_k(i)$ is the kth column of $U(i)$. The integer d is the dimension of M, which will equal the number of non zero (or very small) singular values of $N(i)$ provided that there is sufficient

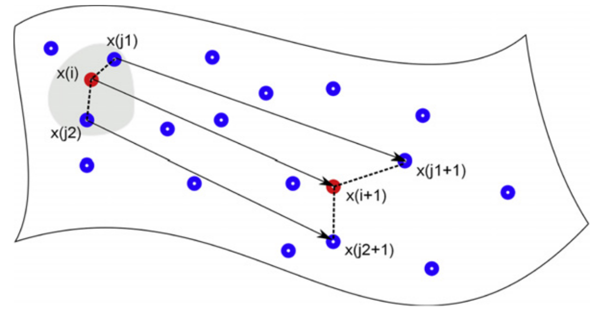

Fig. 6. Reconstruction from data cartoon, grey region represents ball of radius ϵ centred at $x(i)$.

data for a reconstruction—clearly we will need a minimum of d neighbouring points for this to be possible.

The time-τ map takes points close to $\underline{x}(i)$ to points close to $\underline{x}(i+1)$ and we have already shown how to approximate M in these locations. Using these tangent vectors for local co-ordinates at $\underline{x}(i)$ and $\underline{x}(i+1)$ we have

$$\phi_\tau[\underline{x}(i) + \underline{z}] = \underline{x}(i+1) + U(i+1)|_d \overline{N}(i)(V^t(i)S\,|_d^{-1})\underline{z},$$

where

$$U(i+1)|_d = [U_1(i+1)\, U_2(i+1) \cdots U_d(i+1)],$$

$$\overline{N}(i) = \left[\underline{x}(j_1+1) - \underline{x}(i+1)\, \underline{x}(j_2+1) - \underline{x}(i+1) \cdots \underline{x}(j_m+1) - \underline{x}(i+1)\right],$$

$$(V^t(i)S\,|_d^{-1}) = \left[\frac{V_1(i)}{S_{1,1}(i)}\; \frac{V_2(i)}{S_{2,2}(i)} \cdots \frac{V_d(i)}{S_{d,d}(i)}\right].$$

With sufficient data it is possible to reconstruct a system accurately enough to compute its Lyapunov spectrum and even predict its future behaviour.

3.1. Optimal choice of filter for state space reconstructions

We saw in the examples of Section 2 that a smoother kernel results in a smoother 'smoothed' system and that a longer averaging window also gives a less stiff system, as well as better reducing noise. Therefore it is natural to want to use the bump kernel with as long a window as possible.

In the context of state space reconstructions of chaotic systems with smooth and non-smooth non-linearities, where we wish to smooth non-smooth time series data then reconstruct the smoothed system using the method described above, we must balance this improvement in stiffness with a longer window against its negative influence on the reconstruction process. Recall that in order to reconstruct a system from its time series at a point $\underline{x}$ we need to find all of $\underline{x}$'s neighbouring data points, that is all $\underline{y}$ within a distance ϵ of $\underline{x}$. If there are fewer than d such neighbours then it will not be possible to carry out the reconstruction process—the only option being to increase the tolerance ϵ and sacrifice accuracy.

In order for two data points from the smoothed time-series to be within a distance ϵ of each other it is necessary that their T pre images be close together and for their pre images to remain close together backwards in time for the duration of the filtering window. Therefore the longer the filtering window the fewer neighbouring points will be available, without increasing the volume of data or reducing accuracy.

Our suggestion for choosing w is therefore to take the smallest possible value that resolves the non-smooth discontinuities into smooth flow with stiffness equal to the stiffness of the smooth dynamics in the original time-series. Of course, we must also be careful to avoid non-invertibility by comparing the Fourier transform

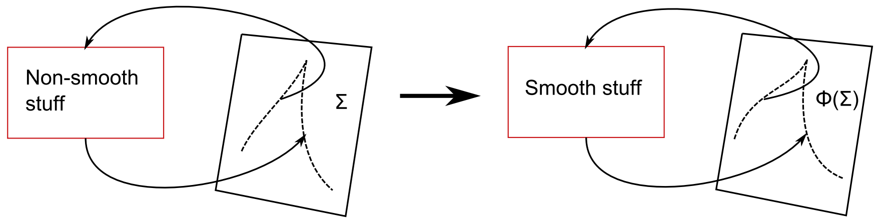

Fig. 7. Poincaré map cartoon.

of the data to the transfer function of the filter as outlined in Section 2.1. This further consideration may fine tune the choice.

The stiffness of a twice differentiable time-series $f : \mathbb{R}_+ \mapsto X$ is given by

$$S[f] = \frac{\max_t \|f''(t)\|}{\max_{t,s} \|f(t) - f(s)\|}.$$

If $f(t)$ is a time series from a non-smooth system then it may only be piecewise twice differentiable, in which case we define the stiffness of its smooth component by taking the maximum of the second derivative only over the twice differentiable pieces

$$S'[f] = \frac{\max_{t: f''(t) < \infty} \|f''(t)\|}{\max_{t,s} \|f(t) - f(s)\|}.$$

4. Smoothness and stability

The easiest way to show that smoothing preserves stability is to consider a Poincaré return map constructed away from any discontinuities. Provided there are no discontinuities in the w-second backward time flow from the section Σ, we can be sure that T is smooth on Σ. The Poincaré maps of the original and smoothed systems are therefore smoothly conjugate and will have exactly the same stability properties (see Fig. 7).

An immediate consequence of this stability equivalence is that our smooth systems will often not be smooth everywhere. If a non-smooth system contains a grazing orbit, say, its return map will contain singular points which will have to be mirrored in the return map of the smoothed system.

Just as we saw in the ad hoc smoothing of the bouncing ball and the moving average smoothing of the graze discontinuity, these singular discontinuities give rise to isolated singularities in the otherwise smooth flow of the transformed system.

5. Numerical example with stability calculation

In this section we will apply our smoothing procedure to time-series data recorded from a computer simulation of a non-linear Duffing impact oscillator. We will then use the smoothed data to calculate the Lyapunov spectrum of the system using a method based on the state space reconstruction technique discussed in Section 3.

The system we use is taken from [8] where Stefanski uses the coupling method to determine the largest Lyapunov exponent of the system from a numerical experiment. This provides us with a standard to test our result against. The system is governed by

$$\ddot{x} = x(1 - x^2) - 0.1\dot{x} + \cos t$$

along with the rule that whenever $\lim_{\tau \to t} x(\tau) = 0.5$, we set $\dot{x}(t) = \lim_{\tau \to t} -0.65\dot{x}(\tau)$. In order to make the system autonomous we include a forcing phase variable θ that obeys $\dot{\theta} = 1$, along with the rule that whenever $\lim_{\tau \to t} \theta(\tau) = 2\pi$, we set $\theta(t) = 0$. This autonomous formulation has two different discontinuities, one associated with the impacting in the oscillator model, and another associated with the phase variable reset. We simulate a long orbit of this chaotic system. The system lives on a strange attractor, which can be broken into three distinct regions depending on which discontinuities points reach forwards and backwards in time. We plot in blue points which have phase reset and are about to impact, green for points which have impacted and are about to impact again and red for points which have impacted and are about to reset, see Fig. 8.

To choose the optimal filter for this problem we use the approach described in Section 3.1. We find that for the bump kernel the window $w = 0.25$ s is the shortest window that resolves the discontinuities into a smooth flow with stiffness less than or equal to that of the smooth non-linearities in the original system, whilst being invertible. See Fig. 9. We apply the bump smoothing transformation to this data. The smoothed data is a single connected component. This transformation is not an injection, so for the calculation we include a delay vector in each of the original variables. Using the state space reconstruction method presented in [7] we compute the Lyapunov spectrum of the system from the smoothed data. We calculate the largest exponent to be 0.0813, which agrees with Stefanski's calculation of 0.0832. Since the system is autonomous and dissipative we know that the second largest eigenvalue is zero and that the third is negative and greater in magnitude than the first. Our results agree with this theory; we have second exponent 0.0031 and third −0.1663. Since our system analyses 6 dimensional data we could produce up to three additional spurious exponents. Our algorithm produces two further finite exponents at much larger order of magnitude and the last exponent becomes infinite during the calculation. See Fig. 10.

6. Discussion

We have shown that low-pass filters can be used to formulate smoothing transformations that map discontinuous or non-differentiable systems to 'smooth' systems—systems which are smooth except for singularities at the images of singular discontinuities such as grazes, cusps and chattering points. We have demonstrated two different techniques for studying these smoothed systems. For simple systems, we can explicitly formulate the transformation and the smoothed system to see how features in the non-smooth flow are integrated into the flow of the smoothed system. For more complicated systems we smooth a time-series, then use state space reconstruction techniques to study the smoothed system.

We have shown that the smoothing procedure preserves stability properties, which gives us a novel way to calculate the Lyapunov spectrum of a non-smooth system. This technique is possible if and only if the smoothing transformation can be chosen to be invertible. In the case that it is invertible then the resulting system still need not be smooth or even have finite Lyapunov exponents, but the Lyapunov exponent calculation will still be correct.

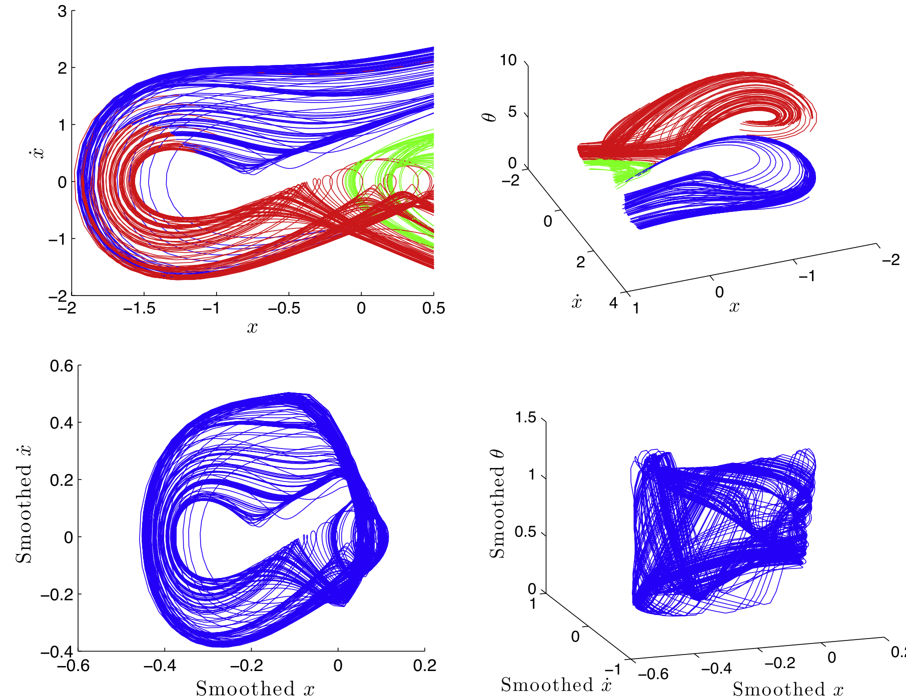

Fig. 8. Top: phase and state space time series for Duffing oscillator, bottom: smoothed phase and state space time series for Duffing oscillator.

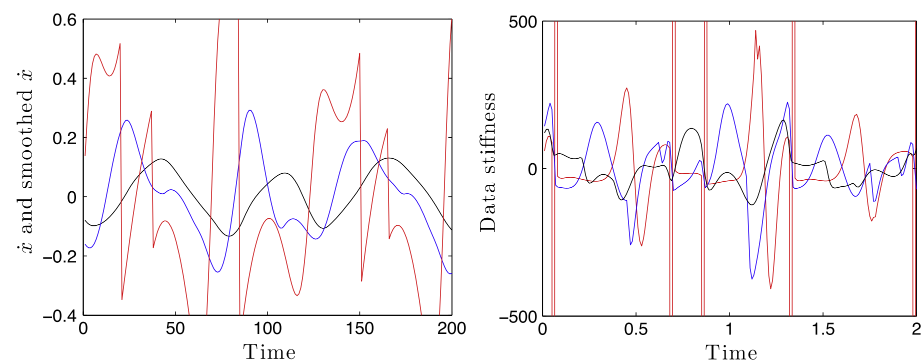

Fig. 9. Left: non-smooth and smoothed time-series, right: stiffness function $f''(t)/\text{range}[f]$. Red: non-smooth time-series, blue: $w = 0.25$ s smoothed time series, black: $w = 0.5$ s smoothed time series. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

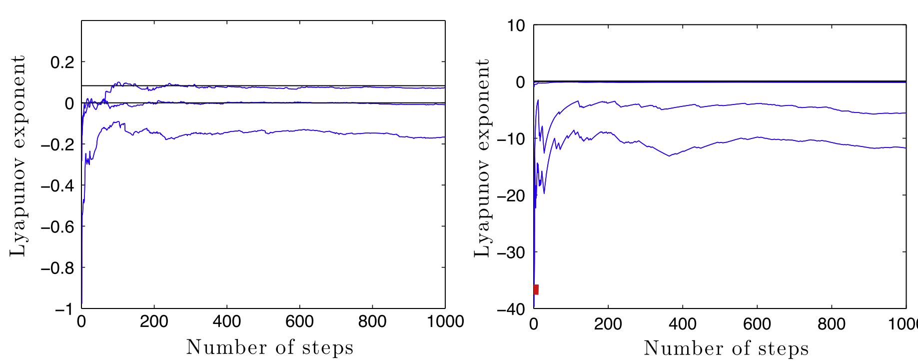

Fig. 10. Left: convergence of three largest Lyapunov exponents, right: behaviour of all 6 exponents, 6th exponent in bold red is taken to be $-\infty$ shortly after experiment begins.

That these smoothing transformations are brought about whenever we apply a low-pass filter to time-series data means that experimentalists may already be inadvertently studying smooth systems of the sort presented here. Our message to experimentalists would be not to avoid this smoothing action by using filters with very short windows or by using more complex smoothing techniques such as the Savitzky Golay smoothing filter. Instead, since it is possible to understand how non-smooth features are transformed by the smoothing, they should apply low-pass filters to reduce the noise then take the effect of the smoothing into account when analysing the data. For example by looking for discontinuities in the second derivative when looking for switches after applying the moving average filter, or looking for square root behaviour in the flow when looking for grazes in a smoothed system.

Some open problems/future work:

- In this paper and also in [5] we derived the smoothed forms for various different discontinuities. But this work is by no means complete. For instance we have not looked at chattering points, cusps, or degenerate grazes at all yet. Can we formulate the smoothed form for all possible discontinuities?
- It is essential that our transformation is invertible. Is it possible to prove some version of Takens' theorem for non-smooth systems that would guarantee the existence of such a transformation?
- What other numerical methods that rely on differentiability or smoothness can we apply to the smoothed data to make inferences about the original non-smooth systems? E.g. techniques for bifurcation continuation?
- What 'real life' applications would benefit from this approach? One possibility is non-smooth electronic circuit dynamics, where analogue sensors are intrinsically low-pass filters.

Acknowledgments

I am very grateful to David Broomhead for many useful discussions and suggestions, particularly in regard to the signal processing side of things. I am also very grateful to Paul Glendinning for his useful comments and suggestions. This work was funded jointly by EPSRC grant EP/E050441/1 "CICADA" and an EPSRC Doctoral Prize Fellowship EP/J50032X/1.

References

[1] A. Champneys, P. Kowalczyk, J. Guckenheimer, M.D. Bernardo, C. Budd, Piecewise-Smooth Dynamical Systems, Springer, 2008.
[2] P. Bloomfield, Fourier Analysis of Time Series, Wiley, 2005.
[3] J. Huke, Embedding nonlinear dynamical systems, a guide to Takens' theorem, MIMS EPrint, Manchester Institute for Mathematical Sciences, University of Manchester, UK, 1997.
[4] B.F. Feeny, Phase–space reconstruction and stick–slip, Nonlinear Dynam. 13 (1997) 39–57.
[5] J. Hook, Topics in dynamical systems, Ph.D. MIMS EPrint, Manchester Institute for Mathematical Sciences, University of Manchester, UK.
[6] A.F. Filippov, Differential Equations with Discontinuous Right-Hand Sides, Springer, 1988.
[7] D. Broomhead, A. Darbyshire, Robust estimation of tangent maps and Liapunov spectra, Physica D 89 (1996) 287–305.
[8] A. Stefanski, Estimation of the largest Lyapunov exponent in systems with impacts, Chaos Solitons Fractals 15 (2000) 2443–2451.