Analysis of Juggling Data

Tolver, Anders; Sørensen, Helle; Muller, Martha; Mousavi, Seyed Nourollah

Published in:
Electronic Journal of Statistics

DOI:
10.1214/14-EJS937F

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Tolver, A., Sørensen, H., Muller, M., & Mousavi, S. N. (2014). Analysis of Juggling Data: Registration of Functional Data Subject to Biomechanical Constraints. Electronic Journal of Statistics, 8(2), 1856-1864. https://doi.org/10.1214/14-EJS937F

Electronic Journal of Statistics
Vol. 8 (2014) 1856–1864
ISSN: 1935-7524
DOI: 10.1214/14-EJS937F

Analysis of juggling data: Registration subject to biomechanical constraints*

Anders Tolver, Helle Sørensen, Martha Muller and Seyed Nourollah Mousavi

Department of Mathematical Sciences
University of Copenhagen
e-mail: tolver@math.ku.dk; helle@math.ku.dk;
m.muller@math.ku.dk; nourollah@math.ku.dk

Abstract: We illustrate how physical constraints of a biomechanical system can be taken into account when registering functional data from juggling trials. We define an idealized model of juggling, based on a periodic joint movement in a low-dimensional space and a periodic position vector (from an undefined joint to the finger tip) of approximately constant length along the observed trajectory. Our registration procedure first warps the cycles in the trial to each other and computes a periodic average, and then estimates the joint movement and the position vector of the abovementioned model.

Keywords and phrases: Biomechanical constraints, decomposition, functional data analysis, juggling trajectories, periodic average, registration, warping.

Received August 2013.

1. Introduction

Functional data are often unsynchronized in their raw form, either due to the sampling process or due to random phase variation (or both). This makes analysis on the raw data problematic since, for example, cross-sectional sample statistics can be misleading. Registration is the process of mapping unsynchronized curves into a synchronized class of functions, with the purpose of effectively filtering out noise before subsequent statistical analyses [1].

At best, registration should use any knowledge of the data generating system, in particular the shape of the underlying signal as well as the nature of possible pertubations. In this paper we discuss registration for functional data from juggling, taking into account simple biomechanical considerations.

Ideally, biomechanics of juggling may be described mathematically by nonlinear dynamical systems, but feedback and feedforward motor control mechanisms are necessary to overrule any disturbed dynamics and thereby impose desired movements or dynamics. We consider data from juggling cycles within in trial as pertubated versions of an idealized periodic movement. The periodic curve represents the average dynamics of the juggling process, whereas the deviations

*Main article 10.1214/14-EJS937.

between the observed data and the idealized signal reflect the complex feedback mechanism between the brain and the motor control system [4].

In conceptualizing an appropriate idealized mathematical model of human juggling, we consider the creation of an electromechanical juggling robot. How would we build and program such a robot? As a minimum, we would construct a rotating finger or hand limb and attach it with a joint to a fixed bar (representing an arm). We could conveniently label the two ends of the hand limb as 'finger tip' and 'joint'.

As a first attempt, we keep the position of the joint fixed and let the position vector from joint to finger tip be periodic. Regarded from a fixed external coordinate frame the position of the finger tip of the robot would trace a trajectory described by

$$f(t) = f_0(t) + c_0$$

where $c_0 \in \mathbb{R}^3$ corresponds to the fixed position of the joint and $f_0 : I \to \mathbb{R}^3$ is the periodic position vector function. Assuming that the robot is a rigid body introduces the geometric constraint that f_0 has constant length, d, such that $|f_0(t)| = d$ for all $t \in I$.

The juggling robot can be improved by allowing the position of the joint to follow a periodic curve. This gives a decomposition of the form

$$f(t) = f_0(t) + c_0(t), \tag{1}$$

where $c_0 : I \to \mathbb{R}^3$ is the trajectory of the joint, while f_0 still describes the vector from joint to finger tip and satisfies $|f_0(t)| = d$ for all $t \in I$ for some d. For identification purposes we assume that c_0 has a simple structure meaning that it belongs to a lower dimensional function space.

In this paper, decompositions of the type (1) will be regarded as idealized juggling signals, and we will demonstrate how to register the observed data towards such idealized signals, i.e. demonstrate that is it is possible to warp and filter the juggling trials such that the resulting curves allow a decomposition of the form (1).

Sections 2 and 3 give a complete description of the registration procedure and details about implementation. In Section 4 we display the results of applying the procedure to the ten trials from the juggling data. Finally, in Section 5 we evaluate the perspectives of combining phase registration and biomechanical constraints.

2. Data and registration procedure

The pre-processed data [2] (lightly smoothed, centered, rotated and trimmed) is the starting point of our analysis, and is referred to as "observed data" or "raw data" in the remainder of the paper. The data indicate the position of the right index finger during juggling and is thus composed of three coordinates. We write $f(t) = (f_1(t), f_2(t), f_3(t))$, and let n denote the number of cycles. There are 10 signals/trials, all collected from the same person. The number of cycles per trial varies from 11 to 13.

The suggested registration procedure is applied to each trial separately, but on all three dimensions and all cycles simultaneously. The implementation details are described in Section 3, but, in short, the complete procedure is split into three steps:

1. **Warping** The observed signal consisting of several cycles is converted into a warped version $f \circ h$, where cycles are warped towards each other using a periodic average function as target for the registration procedure.
2. **Averaging** Based on the warped signal, $f \circ h$, a periodic average, denoted by $\mathcal{P}f$, is computed as a projection onto the (high-dimensional) space of periodic functions.
3. **Decomposition** The periodic average $\mathcal{P}f$ is decomposed into two periodic terms: a joint movement $\mathcal{J}$ belonging to a low-dimensional space, V, and a remainder $\mathcal{P}f - \mathcal{J}f$ with approximately constant length along the trajectory.

The complete procedure involves estimation of a warping function h, a periodic average, and a joint movement $\mathcal{J}f$. Notice that $\mathcal{P}f$ and $\mathcal{J}f$ are periodic per construction, and thus have no between-cycle variation. In particular, we only need to plot the curves on the interval corresponding to one cycle. On the other hand, the warped, but not averaged, curve $f \circ h$ may potentially show amplitude variation between cycles, but presumably only little phase variation, since that has been diminished by warping.

The second step involves projection onto a space of periodic three-dimensional functions. If this projection is denoted by Q_{per}, then $\mathcal{P}f = Q_{per}(f \circ h)$. If $\| \cdot \|$ is the standard L^2-norm and g is a three-dimensional curve, then

$$\frac{\|Q_{per}g\|}{\|g\|} = \sqrt{\frac{\|g\|^2 - \|g - Q_{per}g\|^2}{\|g\|^2}} = \sqrt{1 - \frac{\|g - Q_{per}g\|^2}{\|g\|^2}} \tag{2}$$

takes values in $[0, 1]$ and is a natural measure of the degree of periodicity in g. When data from different cycles are warped against each other as in step 1, we would expect a larger degree of periodicity compared to the raw data. Hence, comparison of $\frac{\|Q_{per}f\|}{\|f\|}$ and $\frac{\|Q_{per}(f \circ h)\|}{\|(f \circ h)\|}$ can be used to quantify the effect of warping on periodicity (see Section 4).

3. Implementation

This section describes technical details of the implementation of our registration procedure. The emphasis is on the decomposition step, since warping and averaging rely on existing techniques and software.

Let f denote a signal consisting of n complete juggling cycles. The duration of each cycle within a trial is rescaled to $[0, 1]$, then the same implementation can be used for all trials, even though the number of cycles are different.

Warping First, we expressed f in terms of 201 Fourier basis functions, and computed the orthogonal projection f_{per} on the space of periodic functions $L_{per,n}$ containing n replications of the same signal. Due to the Fourier basis representation this amounts to keeping coefficients corresponding to harmonics of order $n, 2n, 3n, \ldots, Kn$ (where K is the largest K such that $Kn \leq 100$). Second, a time warping function h maximizing the coherence between $f \circ h$ and f_{per} was estimated. We used the minimal eigenvalue of a cross-product matrix with a roughness penalty on curvature of h as estimation criterion, see [3, Section 7.6]. In order to ensure a sufficient degree of smoothness of the warped signal $f \circ h$ we restricted h to the space spanned by 101 B-splines of order 5 with equally spaced break points. The roughness of the warping functions were controlled by penalizing the squared integral of second order derivatives. The robostness to the value of the penalty parameter λ was examined and for the results presented below we used $\lambda = 10^{-11}$ based on visual inspection.

Averaging The warped function $f \circ h$ was projected onto $L_{per,n}$ (see the paragraph on the warping step above). Hence, we obtain a periodic average of $f \circ h$, denoted $\mathcal{P}f$ and spanned by periodic harmonics.

Decomposition To implement the estimation of $\mathcal{J}f$ in step 3 it was convenient to expand all functions in terms of orthogonal complex exponentials. Denoting by a_k and b_k, $k = 1, 2, 3$, the three coordinate functions of the periodic average $\mathcal{P}f$ (known) and joint movement $\mathcal{J}f$ (to be estimated), we have expansions

$$a_k(t) = \sum_{j=-m}^{m} a_{k,j} \exp(i\omega jt), \quad b_k(t) = \sum_{j=-l}^{l} b_{k,j} \exp(i\omega jt)$$

and hence

$$a_k'(t) = \sum_{j=-m}^{m} i\omega j a_{k,j} \exp(i\omega jt), \quad b_k'(t) = \sum_{j=-l}^{l} i\omega j b_{k,j} \exp(i\omega jt).$$

Here $\omega = 2\pi n$ where n is the number of cycles.

We emphasize that $\mathcal{P}f$ has already been expressed in a finite Fourier basis, thus m and $a_{k,j}$ are all fixed and known at this point of the analysis, whereas the coefficients b_k should be estimated. For $l < m$ fixed, we collect the unknown parameters in θ:

$$\theta = \{b_{k,j} | k = 1, 2, 3, j = -l, \ldots, l\}$$

Some comments on the choice of l: The regularization assumption $l < m$ is necessary for identification, i.e., for the decomposition (1) to be unique since otherwise we could just let $\mathcal{J}f = \mathcal{P}f - c_0$ with $c_0 \in \mathbb{R}^3$ any fixed vector. For $l < m$ the joint movement $\mathcal{J}f$ belongs to a subspace of lower dimension than $\mathcal{P}f$, and the idea is to choose a small l, such that the joint movement is simple.

Recall that we aim at finding $\mathcal{J}f$ such that $\mathcal{P}f - \mathcal{J}f$ has approximately constant length; hence we want the derivative of the squared length to be approximately zero for all t:

$$D|\mathcal{P}f(t) - \mathcal{J}f(t)|^2 \approx 0.$$

This leads to the following criterion function to be minimized:

$$
\begin{aligned}
C(\theta) &= \int_0^1 \left[D|\mathcal{P}f(t) - \mathcal{J}f(t)|^2\right]^2 dt \qquad (3) \\
&= \int_0^1 \left[D \sum_{k=1}^{3} (a_k(t) - b_k(t))^2\right]^2 dt \\
&= 4 \int_0^1 \left[\sum_{k=1}^{3} D(a_k(t) - b_k(t)) \cdot (a_k(t) - b_k(t))\right]^2 dt.
\end{aligned}
$$

If we introduce the notation $e_{k,j} = a_{k,j} - b_{k,j}$ (with $b_{k,j} = 0, |k| > l$) for the Fourier coefficients of the difference $\mathcal{P}f - \mathcal{J}f$, and furthermore $c_{j_1,j_2} = \{\sum_{k=1}^{3} j_2 e_{k,j_1} e_{k,j_2}\}$ and let $j \in I_s$ if $j, s-j \in \{-m, \ldots, m\}$, then

$$
C(\theta) = \int_0^1 \left[\sum_{s=-2m}^{2m} i\omega \sum_{j \in I_s} c_{s-j,j} \exp(i\omega st)\right]^2 dt.
$$

Finally, if we let $d_s = \sum_{j \in I_s} c_{s-j,j}$ and use that $d_{-s} = -\overline{d_s}$ (complex conjugate), then we end up with the following simple formula for the criterion function

$$
C(\theta) = -4\omega^2 \sum_{s=-2m}^{2m} d_s d_{-s} = 4\omega^2 \left\{|d_0|^2 + 2\sum_{s=1}^{2m} |d_s|^2\right\}. \qquad (4)
$$

The representation (4) makes it feasible to compute numerically the value and the gradient of the objective function as a function of θ to be used for the minimization algorithm. Since we are looking for a real valued estimate of the joint movement $\mathcal{J}f$, we found it convenient to reparameterize the problem in terms of a basis of sines and cosines. For the results below we used $l = 1$ corresponding to the joint movement being expressed in terms of first order harmonics only.

4. Results

We applied the registration procedure described above to each of the ten juggling trials. We will use trial 8 for detailed illustration, because the effect of the warping step was largest for this trial.

Warping and averaging Figure 1 shows the effect of steps 1 and 2 (warping and averaging) on trial 8. The vertical coordinate (z)of the raw data (dashed) is shown together with vertical coordinate of the periodic signal $\mathcal{P}f$ (solid). The raw signal does not exhibit much misalignment but the signal is indeed warped slightly. Notice how the warping is more pronounced towards the ends of the trial. The average curve $\mathcal{P}f$ for trial 8 is shown for each coordinate separately in the left part of Figure 2, and as a 3d-curve in the right part of the figure (solid curve).

For the raw data the degree of periodicity, cf. definition (2), was 88.0%, whereas for the warped data this number increased to 98.6%. All other trials had degrees of periodicity of 94.3% to 97.2% before warping and between 97.5%

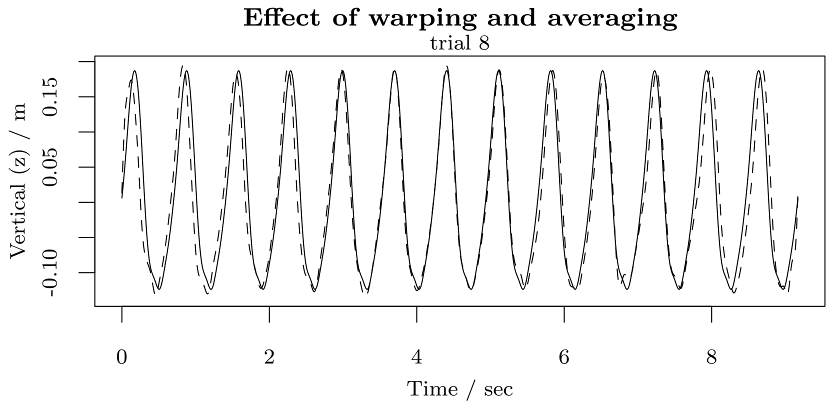

FIG 1. *Warping and averaging for trial 8. The dashed curve shows the z coordinate of the observed data, while the solid curve shows the z coordinate of $\mathcal{P}f$.*

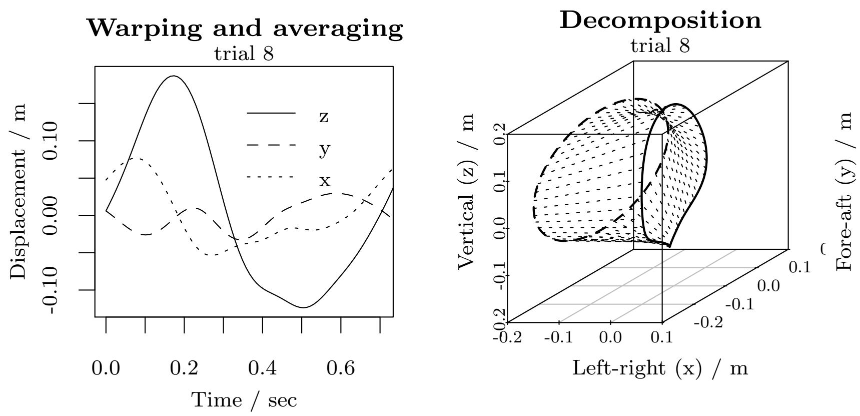

FIG 2. *Left: The three directions of the warped and averaged curve $\mathcal{P}f$ for trial 8. Right: 3d-illustration of the decomposition for trial 8. The solid curve shows the average $\mathcal{P}f$, the dashed curve shows the estimated joint movement curve $\mathcal{J}f$, and the dotted lines illustrate the trajectory of the difference $\mathcal{P}f - \mathcal{J}f$ (each dotted line correspond to a specific time point.)*

and 99.2% after warping. Hence, in general, only a limited amount of warping towards the periodic template was necessary. Visually, the raw and averaged trials were almost indistinguishable, except for trial 8 (see Figure 1).

The upper left, upper right and lower left plots of Figure 3 show the three coordinates of the warped curves $f \circ h$ for all ten trials, split into cycles and rescaled to the unit interval. The curves are coloured according to trial (but note that curves from different trials have not been aligned). In general, cycles within a trial are well aligned. Therefore the projection onto $L_{per,n}$ is a good representation of a trial. Note that the projections are similar across trials (-see the lower right part of Figure 3). The warping criterion gives less weight to coordinates with lower amplitude variation. This may explain why most misalignment is present in the y direction.

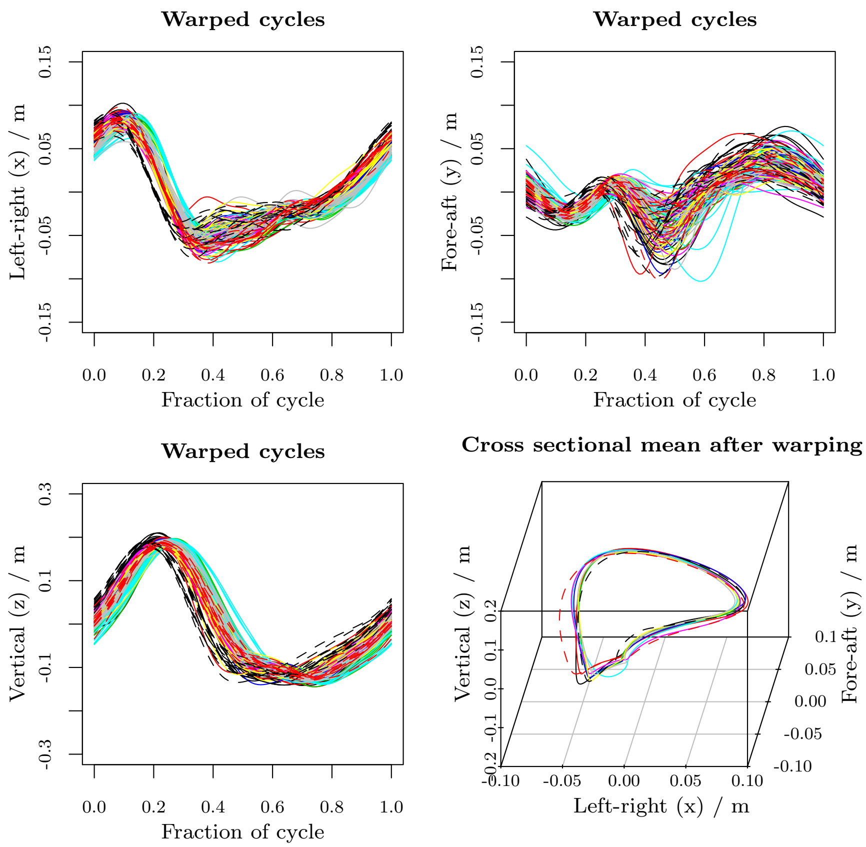

FIG 3. *Upper left, upper right and lower left: The three coordinates of the warped curves $f \circ h$ cut into individual cycles for each trial. For a trial with n cyc les, the complete curve was simply divided into n pieces of the same length, which was then rescaled to the unit interval. Cycles of the same colour and line type stem from the same trial. Lower right: 3d-scatterplot of the periodic average $\mathcal{P}f$ for all trials.*

Decomposition The estimated joint movement $\mathcal{J}f$ for trial 8 is shown as a dashed curve in the right part of Figure 2. Recall that the estimation procedure seeks the curve $\mathcal{J}f$ such that the vector $\mathcal{P}f - \mathcal{J}f$ has approximately constant length over the trajectory. This vector is illustrated by the dotted lines between the two curves, and its length varies from 0.179 m to 0.182 m for trial 8.

The decompositions for all curves are illustrated in Figure 4. The left part shows the length $|\mathcal{P}f - \mathcal{J}f|$ over the trajectories (scaled to the unit interval), and the right part shows the joint movements $\mathcal{J}f$. We make the following immediate observations from Figure 4: First, for all ten trials it was possible to obtain a function $\mathcal{J}f \in V$ such that the distance $|\mathcal{P}f - \mathcal{J}f|$ is approximately constant over time. This indicates that our simplistic biomechanical considerations leading to equation (1) characterizes some of the main features of the data generating mechanism. Second, the estimated length varies from 0.077 m

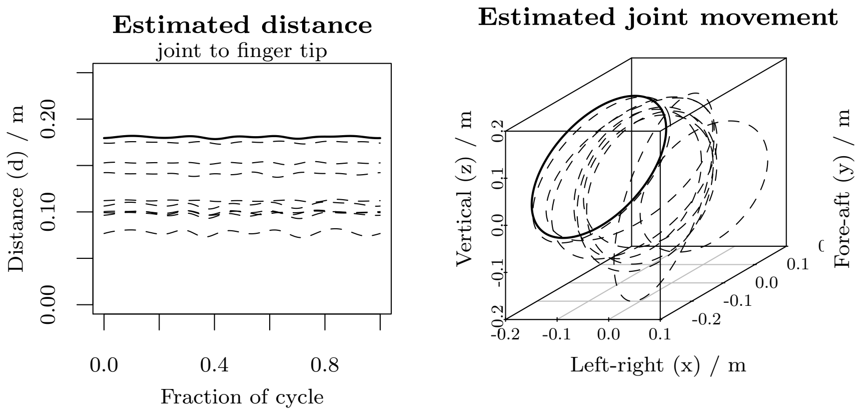

FIG 4. *Left: Estimated trajectory of distances, $|\mathcal{P}f - \mathcal{J}f|$, for all 10 trials. Right: Estimated joint movement, $\mathcal{J}f$, for all ten trials. In both plots the estimate correponding to trial 8 is shown as a solid curve.*

to 0.181 m across the ten trials. This is somewhat disappointing as we had hoped for an interpretation of this length as the length of a part of the hand or arm of the juggler. Third, the variation between the estimated joint movement curves is substantial. The decomposition restricts $\mathcal{J}f$ to be spanned by first order harmonics in all three directions. Allthough the curves are approximately elliptic they are different regarding angle and position.

5. Discussion

The purpose of the paper was to illustrate how the physical nature of a biomechanical system could be taken into account when removing phase variation of functional data from juggling. We have demonstrated that it is possible to warp all ten juggling trials such that the resulting structural mean over all cycles allows a decomposition as in (1).

The most striking observation is that the estimated distance from finger tip to joint, which should be an internal constant of the body anatomy, varies substantially across the ten trials. This complicates the physical interpretation of the estimated decomposition. Looking more carefully at the curves in the left part of Figure 4, there seems to be some common patterns in the deviations from constancy. Curves with low values of d seem to have peaks and valleys at the same time points (for example around 0.38 and 0.82), i.e. at the same time points of the juggling cycle. This indicates that our simple model might not have captured all features in the data.

A possible extension of the model would be to allow for more flexibility in the space V for the joint movement, i.e. by introducing harmonics of higher order in the basis for $\mathcal{J}f$. However, it seems more likely that adjustments from the idealized set-up given by (1) is taking place around the finger tip (far from the corpus) rather than at joints closer to the corpus. This suggest to relax the focus on constant length of $\mathcal{P}f - \mathcal{J}f$. For example, the criterion function $C(\theta)$ in the decomposition step, see (3) and (4), could be adjusted to have a time-varying penalty on deviations from constancy. This would, however, complicate the optimization problem substantially.

In this connection, it should be mentioned that the numerical optimization problem for estimating the decomposition was more challenging than expected. The algorithm we used produced reliable estimates but was slow. This part of the implementation could be improved.

It is important to realize that amplitude and phase variation are bound to be intertwined, as an adjustment via a change in speed (phase) will most likely also change the amplitude. In relation to this, the complicated interplay between the estimation the warping function (step 1) and the estimation of the joint movement (step 3) should also be noticed. In particular, the space V for the joint movement is not invariant to warping (i.e. $g \in V$ does not imply that $g \circ h \in V$ for a warping function h). Too much warping of f may destroy the interpretation of the decomposition. This could be avoided by simultaneously estimating the warping function and the decomposition, i.e. to incorporate the warping (and averaging) step into the decomposition step.

Apart from the suggestions mentioned above, it would be interesting to examine the robustness of the registration. Simulations could clarify the importance of the explicit form of the underlying signal on the performance of the registration procedure. Moreover, it would be interesting to fit a common joint movement curve to all s, and see the effect on the corresponding position vectors $\mathcal{P}f - \mathcal{J}f$ and their lengths.

Acknowledgements

We acknowledge the Mathematical Biosciences Institute, Ohio, for supporting our participation in the workshop on "Statistics of Time Warpings and Phase Variations".

References

[1] Kneip, A. and Ramsay, J. O. (2008). Combining registration and fitting for functional models. *Journal of the American Statistical Association* **103**, 483, 1155–1165. http://amstat.tandfonline.com/doi/abs/10.1198/016214508000000517. MR2528838

[2] Ramsay, J. O., Gribble, P., and Kurtek, S. (2014). Description and processing of functional data arising from juggling trajectories. *Electron. J. Statist.* **8**, 1811–1816, Special Section on Statistics of Time Warpings and Phase Variations.

[3] Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*, Second ed. Springer, New York. MR2168993

[4] Schaal, S., Atkeson, C. G., and Sternad, D. (1996). One-handed juggling: A dynamical approach to a rhythmic movement task. *Journal of Motor Behavior* **28**, 2, 165–183.