A Variational Model for Data Fitting on Manifolds by Minimizing the Acceleration of a Bézier Curve

Ronny Bergmann[a], Pierre-Yves Gousenbourger[b]

[a]Technische Universität Chemnitz, Chemnitz, Germany

[b]Université catholique de Louvain, Louvain-la-Neuve, Belgium.

Minisymposium MS FT-1-SG 4: Manifold Sensing and Sparse Recovery,

International Congress on Industrial and Applied Mathematics,

Valencia, Spain,

July 16, 2019.

Data Fitting on Manifolds

Given data points $d_0, \ldots, d_n$ on a Riemannian manifold $\mathcal{M}$ and time points $t_i \in I$, find a "nice" curve $\gamma\colon I \to \mathcal{M}$, $\gamma \in \Gamma$, such that $\gamma(t_i) = d_i$ (interpolation) or $\gamma(t_i) \approx d_i$ (approximation).

Data Fitting on Manifolds

Given data points $d_0, \dots, d_n$ on a Riemannian manifold $\mathcal{M}$ and time points $t_i \in I$, find a "nice" curve $\gamma\colon I \to \mathcal{M}$, $\gamma \in \Gamma$, such that $\gamma(t_i) = d_i$ (interpolation) or $\gamma(t_i) \approx d_i$ (approximation).

- Γ set of geodesics & approximation: geodesic regression
  [Rentmeesters, 2011; Fletcher, 2013; Boumal, Absil, 2011]
- Γ Sobolev space of curves: Inifinite-dimensional problem
  [Samir et al., 2012]
- Γ composite Bézier curves; LSs in tangent spaces
  [Arnould et al., 2015; Gousenbourger, Massart, Absil, 2018]
- Discretized curve, $\Gamma = \mathcal{M}^N$, [Boumal, Absil, 2011]

Data Fitting on Manifolds

Given data points $d_0, \dots, d_n$ on a Riemannian manifold $\mathcal{M}$ and time points $t_i \in I$, find a "nice" curve $\gamma\colon I \to \mathcal{M}$, $\gamma \in \Gamma$, such that $\gamma(t_i) = d_i$ (interpolation) or $\gamma(t_i) \approx d_i$ (approximation).

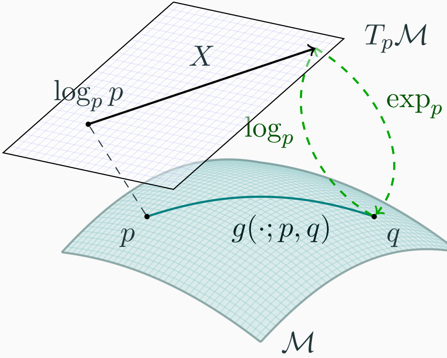
- Γ set of geodesics & approximation: geodesic regression [Rentmeesters, 2011; Fletcher, 2013; Boumal, Absil, 2011]
- Γ Sobolev space of curves: Inifinite-dimensional problem [Samir et al., 2012]
- Γ composite Bézier curves; LSs in tangent spaces [Arnould et al., 2015; Gousenbourger, Massart, Absil, 2018]
- Discretized curve, $\Gamma = \mathcal{M}^N$, [Boumal, Absil, 2011]

This talk.

"nice" means minimal (discretized) acceleration ("as straight as possible") for Γ the set of composite Bézier curves.

Closed form solution for $\mathcal{M} = \mathbb{R}^d$: Natural (cubic) splines.

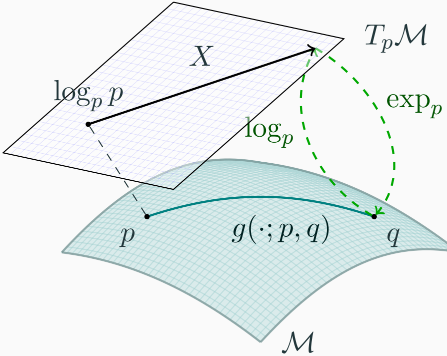
A d-dimensional Riemannian Manifold $\mathcal{M}$



A d-dimensional Riemannian manifold can be informally defined as a set $\mathcal{M}$ covered with a 'suitable' collection of charts, that identify subsets of $\mathcal{M}$ with open subsets of $\mathbb{R}^d$ and a continuously varying inner product on the tangential spaces.

[Absil, Mahony, Sepulchre, 2008]

A d-dimensional Riemannian Manifold $\mathcal{M}$



Geodesic $g(\cdot; p, q)$ shortest path (on $\mathcal{M}$) between $p, q \in \mathcal{M}$

Tangent space $\mathrm{T}_p\mathcal{M}$ at p, with inner product $(\cdot, \cdot)_p$

Logarithmic map $\log_p q = \dot{g}(0; p, q)$ "speed towards q"

Exponential map $\exp_p X = g(1)$, where $g(0) = p$, $\dot{g}(0) = X$

Variational Methods on Manifolds

Variational methods model a trade-off between staying close to the data and minimizing a certain property

$$\mathcal{E}(p) = D(p; f) + \alpha R(p), \quad p \in \mathcal{M}$$

- $\alpha > 0$ is a weight
- $\mathcal{M}$ is a Riemannian manifold
- given (input) data $f \in \mathcal{M}$
- data or similarity term $D(p; f)$
- regularizer / prior $R(p)$
- $\mathcal{E}$ is smooth, but high-dimensional, $\mathcal{M} = \mathcal{N}^m$, $m \in \mathbb{N}$

(Euclidean) Bézier Curves

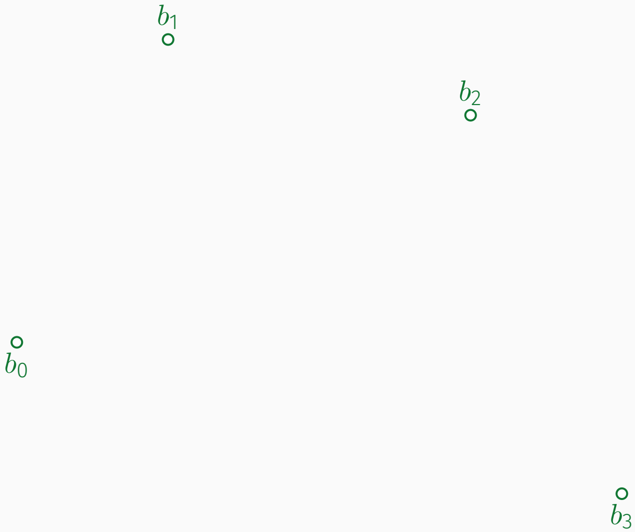
Definition [Bézier, 1962]
A Bézier curve β_K of degree $K \in \mathbb{N}_0$ is a function $\beta_K\colon [0,1] \to \mathbb{R}^d$ parametrized by control points $b_0, \ldots, b_K \in \mathbb{R}^d$ and defined by

$$\beta_K(t; b_0, \ldots, b_K) := \sum_{j=0}^{K} b_j B_{j,K}(t),$$

[Bernstein, 1912]
where $B_{j,K} = \binom{K}{j} t^j (1-t)^{K-j}$ are the Bernstein polynomials of degree K.

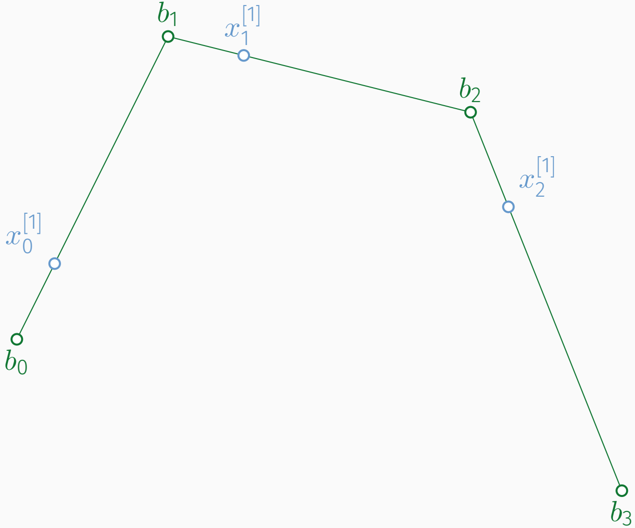
Evaluation via Casteljau's algorithm. [de Casteljau, 1959]

Illustration of de Casteljau's Algorithm



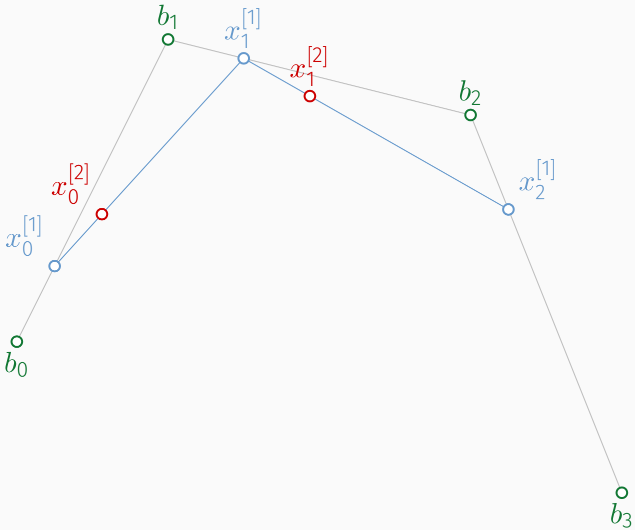
The set of control points b_0, b_1, b_2, b_3.

Illustration of de Casteljau's Algorithm



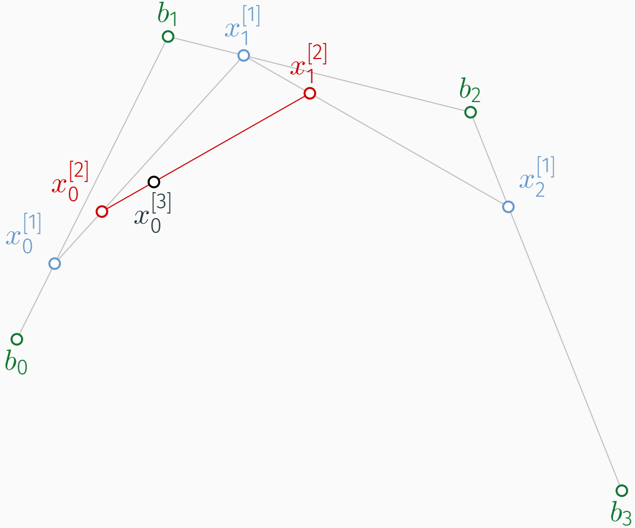
Evaluate line segments at $t = \frac{1}{4}$, obtain $x_0^{[1]}, x_1^{[1]}, x_2^{[1]}$.

Illustration of de Casteljau's Algorithm



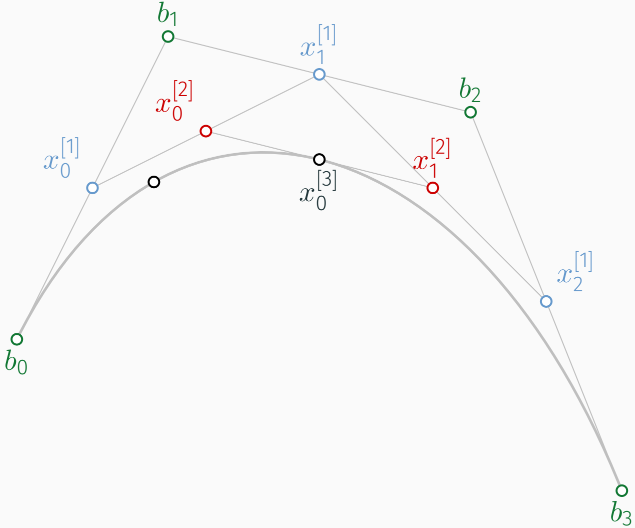
Repeat evaluation for new line segments to obtain $x_0^{[2]}, x_1^{[2]}$.

Illustration of de Casteljau's Algorithm



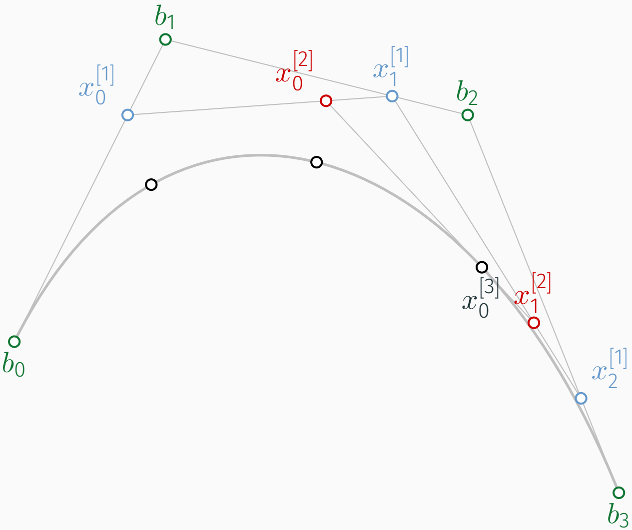
Repeat for the last segment to obtain $\beta_3(\frac{1}{4}; b_0, b_1, b_2, b_3) = x_0^{[3]}$.

Illustration of de Casteljau's Algorithm



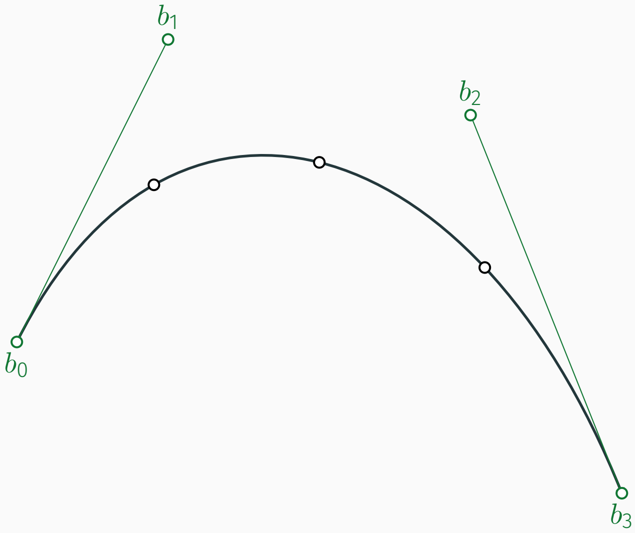
Same procedure for evaluation of $\beta_3(\frac{1}{2}; b_0, b_1, b_2, b_3)$.

Illustration of de Casteljau's Algorithm



Same procedure for evaluation of $\beta_3(\frac{3}{4}; b_0, b_1, b_2, b_3)$.

Illustration of de Casteljau's Algorithm



Complete curve $\beta_3(t; b_0, b_1, b_2, b_3)$.

Composite Bézier Curves

Definition
A composite Bezier curve $B\colon [0,n] \to \mathbb{R}^d$ is defined as

$$B(t) := \begin{cases} \beta_K(t; b_0^0, \dots, b_K^0) & \text{if } t \in [0,1], \\ \beta_K(t-i; b_0^i, \dots, b_K^i), & \text{if } t \in (i, i+1], \quad i = 1, \dots, n-1. \end{cases}$$

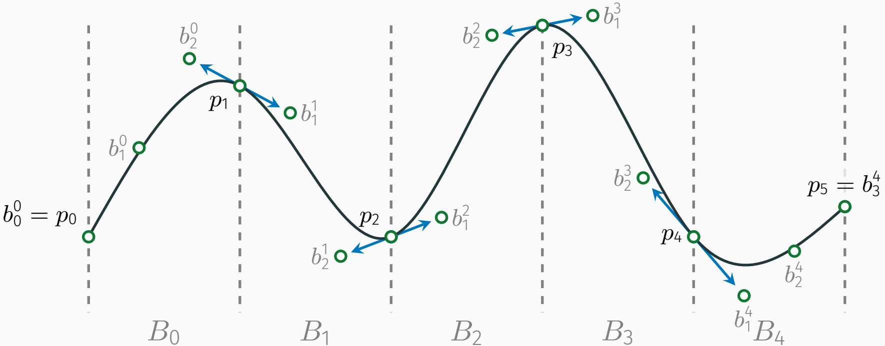
Composite Bézier Curves

Definition

A composite Bezier curve $B\colon [0,n] \to \mathbb{R}^d$ is defined as

$$B(t) := \begin{cases} \beta_K(t; b_0^0, \dots, b_K^0) & \text{if } t \in [0,1], \\ \beta_K(t-i; b_0^i, \dots, b_K^i), & \text{if } t \in (i, i+1], \quad i = 1, \dots, n-1. \end{cases}$$

Denote ith segment by $B_i(t) = \beta_K(t; b_0^i, \dots, b_K^i)$ and $p_i = b_0^i$.



Composite Bézier Curves

Definition
A composite Bezier curve $B\colon [0,n] \to \mathbb{R}^d$ is defined as

$$B(t) := \begin{cases} \beta_K(t; b_0^0, \dots, b_K^0) & \text{if } t \in [0,1], \\ \beta_K(t-i; b_0^i, \dots, b_K^i), & \text{if } t \in (i, i+1], \quad i = 1, \dots, n-1. \end{cases}$$

Denote ith segment by $B_i(t) = \beta_K(t; b_0^i, \dots, b_K^i)$ and $p_i = b_0^i$.

- continuous iff $B_{i-1}(1) = B_i(0)$, $i = 1, \dots, n-1$
  $\Rightarrow b_K^{i-1} = b_0^i = p_i$, $i = 1, \dots, n-1$
- continuously differentiable iff $p_i = \frac{1}{2}(b_{K-1}^{i-1} + b_1^i)$

Definition. [Park, Ravani, 1995; Popiel, Noakes, 2007]
Let $\mathcal{M}$ be a Riemannian manifold and $b_0, \ldots, b_K \in \mathcal{M}$, $K \in \mathbb{N}$.

The (generalized) Bézier curve of degree k, $k \leq K$, is defined as

$$\beta_k(t; b_0, \ldots, b_k) = g(t; \beta_{k-1}(t; b_0, \ldots, b_{k-1}), \beta_{k-1}(t; b_1, \ldots, b_k)),$$

if $k > 0$, and

$$\beta_0(t; b_0) = b_0.$$

Definition. [Park, Ravani, 1995; Popiel, Noakes, 2007]
Let $\mathcal{M}$ be a Riemannian manifold and $b_0, \dots, b_K \in \mathcal{M}$, $K \in \mathbb{N}$.

The (generalized) Bézier curve of degree k, $k \leq K$, is defined as

$$\beta_k(t; b_0, \dots, b_k) = g(t; \beta_{k-1}(t; b_0, \dots, b_{k-1}), \beta_{k-1}(t; b_1, \dots, b_k)),$$

if $k > 0$, and

$$\beta_0(t; b_0) = b_0.$$

- Bézier curves $\beta_1(t; b_0, b_1) = g(t; b_0, b_1)$ are geodesics.
- composite Bézier curves $B\colon [0, n] \to \mathcal{M}$ completely analogue (using geodesics for line segments)

Definition. [Park, Ravani, 1995; Popiel, Noakes, 2007]
Let $\mathcal{M}$ be a Riemannian manifold and $b_0, \ldots, b_K \in \mathcal{M}$, $K \in \mathbb{N}$.

The (generalized) Bézier curve of degree k, $k \leq K$, is defined as

$$\beta_k(t; b_0, \ldots, b_k) = g(t; \beta_{k-1}(t; b_0, \ldots, b_{k-1}), \beta_{k-1}(t; b_1, \ldots, b_k)),$$

if $k > 0$, and

$$\beta_0(t; b_0) = b_0.$$

The Riemannian composite Bezier curve $B(t)$ is

- continuous iff $B_{i-1}(1) = B_i(0)$, $i = 1, \ldots, n-1$
  $\Rightarrow b_K^{i-1} = b_0^i =: p_i$, $i = 1, \ldots, n-1$
- continuously differentiable iff $p_i = g\big(\frac{1}{2}; b_{K-1}^{i-1}, b_1^i\big)$

Definition. [Park, Ravani, 1995; Popiel, Noakes, 2007]
Let $\mathcal{M}$ be a Riemannian manifold and $b_0, \dots, b_K \in \mathcal{M}$, $K \in \mathbb{N}$.

The (generalized) Bézier curve of degree k, $k \leq K$, is defined as

$$\beta_k(t; b_0, \dots, b_k) = g(t; \beta_{k-1}(t; b_0, \dots, b_{k-1}), \beta_{k-1}(t; b_1, \dots, b_k)),$$

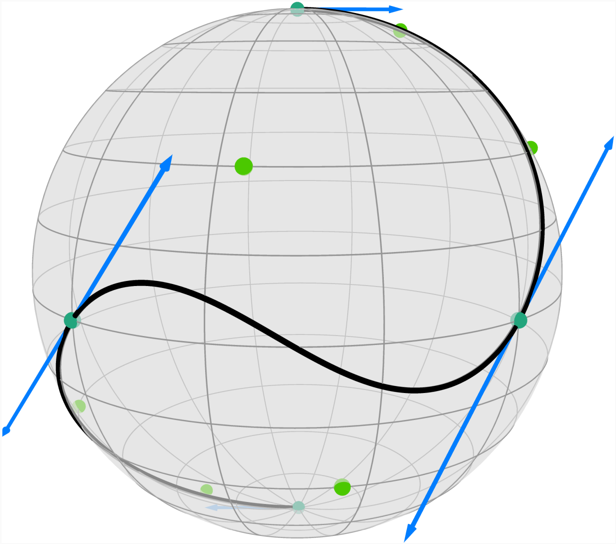
if $k > 0$, and

$$\beta_0(t; b_0) = b_0.$$

The Riemannian composite Bezier curve $B(t)$ is

- continuous iff $B_{i-1}(1) = B_i(0)$, $i = 1, \dots, n-1$
  $\Rightarrow b_K^{i-1} = b_0^i =: p_i$, $i = 1, \dots, n-1$
- continuously differentiable iff $b_{K-1}^{i-1} = g\big(2; b_1^i, p_i\big)$.

Illustration of a Composite Bézier Curve on the Sphere $\mathbb{S}^2$



The directions, e.g. $\log_{p_j} b_j^1$, are now tangent vectors.

A Variational Model for Data Fitting

Let $d_0, \ldots, d_n \in \mathcal{M}$. A model for data fitting reads

$$\mathcal{E}(B) = \frac{\lambda}{2} \sum_{k=0}^{n} d^2_{\mathcal{M}}(B(k), d_k) + \int_0^n \Big\| \frac{D^2 B(t)}{\mathrm{d}t^2} \Big\|^2_{B(t)} \mathrm{d}t, \qquad \lambda > 0,$$

where $B \in \Gamma$ is from the set of continuously differentiable composite Bezier curve of degree K with n segments.

A Variational Model for Data Fitting

Let $d_0, \ldots, d_n \in \mathcal{M}$. A model for data fitting reads

$$\mathcal{E}(B) = \frac{\lambda}{2} \sum_{k=0}^{n} d^2_{\mathcal{M}}(B(k), d_k) + \int_0^n \Bigl\lVert \frac{D^2 B(t)}{\mathrm{d}t^2} \Bigr\rVert^2_{B(t)} \mathrm{d}t, \qquad \lambda > 0,$$

where $B \in \Gamma$ is from the set of continuously differentiable composite Bezier curve of degree K with n segments.

- Goal: find minimizer $B^* \in \Gamma$
- finite dimensional optimization problem in the control points b^i_j, i.e. on $\mathcal{M}^L$ with
  - $L = n(K-1) + 2$
  - $\lambda \to \infty$ yields interpolation $(p_k = d_k) \Rightarrow L = n(K-2) + 1$

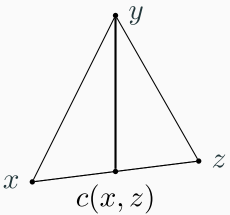
Second Order Differences on Manifolds

Second order difference: [RB et al., 2014; RB, Weinmann, 2016; Bačák et al., 2016]

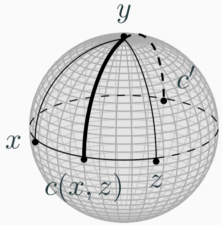
$$\mathrm{d}_2(x,y,z) \coloneqq \min_{c\in\mathcal{C}_{x,z}} \mathrm{d}_{\mathcal{M}}(c,y), \quad x,y,z\in\mathcal{M},$$

$\mathcal{C}_{x,z}$ mid point(s) of geodesic(s) $g(\cdot;x,z)$

$\frac{1}{2}\|x-2y+z\|_2 = \|\frac{1}{2}(x+z)-y\|_2$



$\min_{c\in\mathcal{C}_{x,z}} d_{\mathcal{M}}(c,y)$



$\mathcal{M}=\mathbb{S}^2$

Discretizing the Data Fitting Model

We discretize the absolute second order covariant derivative

$$\int_0^n \left\| \frac{D^2 B(t)}{\mathrm{d}t^2} \right\|_{\gamma(t)}^2 \mathrm{d}t \approx \sum_{k=1}^{N-1} \frac{\Delta_s d_2^2[B(s_{i-1}), B(s_i), B(s_{i+1})]}{\Delta_s^4}.$$

on equidistant points $s_0, \ldots, s_N$ with step size $\Delta_s = s_1 - s_0$.

Evaluating $\mathcal{E}(B)$ consists of evaluation of geodesics and squared (Riemannian) distances

- $(N+1)K$ geodesics to evaluate the Bézier segments
- N geodesics to evaluate the mid points c
- N squared distances to obtain the second order absolute finite differences squared

Gradient of the Discretized Data Fitting Model

For the gradient of the discretized data fitting model

$$\mathcal{E}(B) = \frac{\lambda}{2}\sum_{k=0}^{n} d^2_{\mathcal{M}}(B(k), d_k) + \sum_{k=1}^{N-1} \frac{\Delta_s d_2^2[B(s_{i-1}), B(s_i), B(s_{i+1})]}{\Delta_s^4}.$$

we

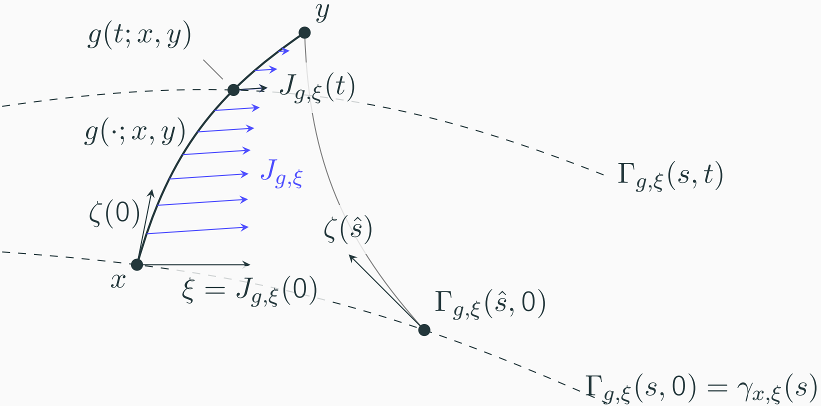
- identified first and last control points $p_i = b^{i-1}_K = b^i_0$
- plug in the constraint $b^{i-1}_{K-1} = g(2; b^i_1, p_i)$
  $\Rightarrow$ Introduces a further chain rule for the differential
  $\Rightarrow$ reduces the number of optimization variables.
- concatenation of adjoint Jacobi fields (evaluated at the points s_i) yields the gradient $\nabla_{\mathcal{N}}\mathcal{E}$.

The Differential of a Geodesic w.r.t. its Start Point

The geodesic variation

$$\Gamma_{g,\xi}(s,t) := \exp_{\gamma_{x,\xi}(s)}(t\zeta(s)), \qquad s \in (-\varepsilon,\varepsilon),\ t \in [0,1], \varepsilon > 0.$$

is used to define the Jacobi field $J_{g,\xi}(t) = \frac{\partial}{\partial s}\Gamma_{g,\xi}(s,t)|_{s=0}$.



Then the differential reads $D_x g(t;\cdot,y)[\xi] = J_{g,\xi}(t)$.

Implementing Jacobi Fields

- On symmetric manifolds, the Jacobi field can be evaluated in closed form, since the PDE decouples into ODEs.
- The adjoint Jacobi fields $J^*_{g,\eta}(t)$ are characterized by

$$\langle J_{g,\xi}(t), \eta\rangle_{g(t)} = \langle \xi, J^*_{g,\eta}(t)\rangle_x, \quad \text{for all } \xi \in T_x\mathcal{M}, \eta \in T_{g(t;x,y)}\mathcal{M}$$

can be computed without extra efforts, i.e. the same ODEs occur.

$\Rightarrow$ adjoint Jacobi fields can be used to calculate the gradient

- Gradient of iterated evaluations of geodesics can be computed by composition of (adjoint) Jacobi fields

Gradient Descent on a Manifold

Let $\mathcal{N} = \mathcal{M}^L$ be the product manifold of $\mathcal{M}$,

Input.

- $\mathcal{E} \colon \mathcal{N} \to \mathbb{R}$,
- its gradient $\nabla_{\mathcal{N}}\mathcal{E}$,
- initial data $q^{(0)} = b \in \mathcal{N}$
- step sizes $s_k > 0, k \in \mathbb{N}$.

Output: $\hat{q} \in \mathcal{N}$

$k \leftarrow 0$

repeat

$\quad q^{(k+1)} \leftarrow \exp_{q^{(k)}}\bigl(-s_k \nabla_{\mathcal{N}}\mathcal{E}(q^{(k)})\bigr)$

$\quad k \leftarrow k + 1$

until a stopping criterion is reached

return $\hat{q} \coloneqq q^{(k)}$

Armijo Step Size Rule

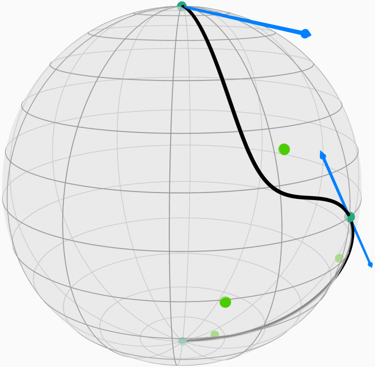
Let $q = q^{(k)}$ be an iterate from the gradient descent algorithm, $\beta, \sigma \in (0, 1), \alpha > 0$.

Let m be the smallest positive integer such that

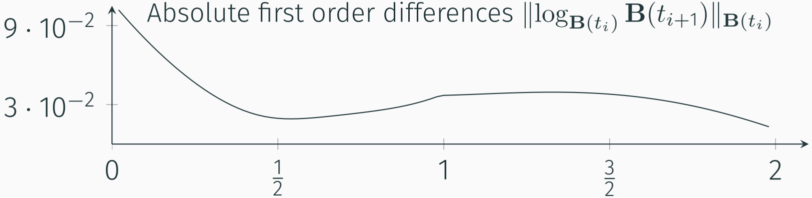
$$\mathcal{E}(q) - \mathcal{E}\big(\exp_q(-\beta^m \alpha \nabla_{\mathcal{N}} \mathcal{E}(q))\big) \geq \sigma \beta^m \alpha \|\nabla_{\mathcal{N}} \mathcal{E}(q)\|_q$$

holds. Set the step size $s_k \coloneqq \beta^m \alpha$.

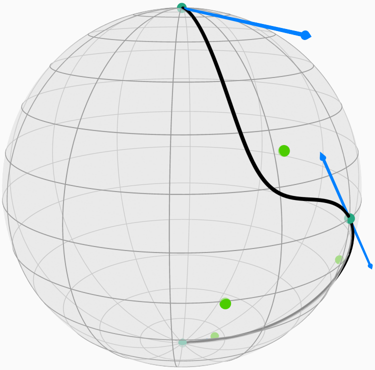
Minimizing with Known Minimizer



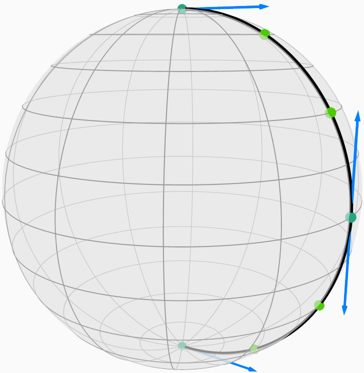
Original



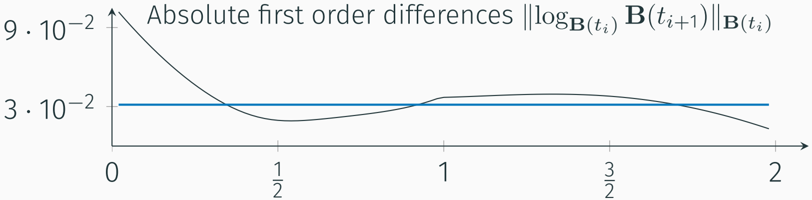
Minimizing with Known Minimizer



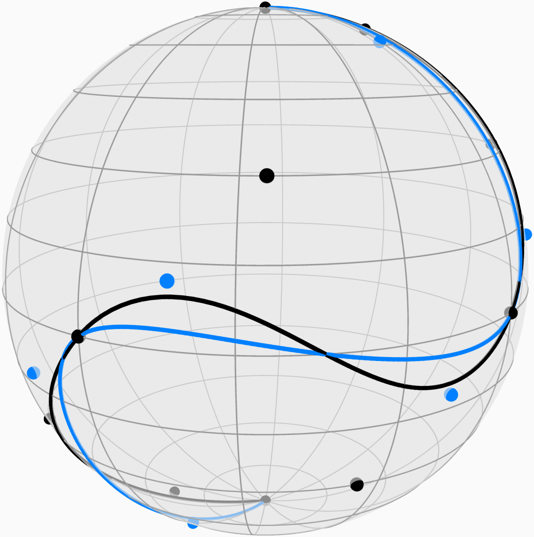
Original



Minimized



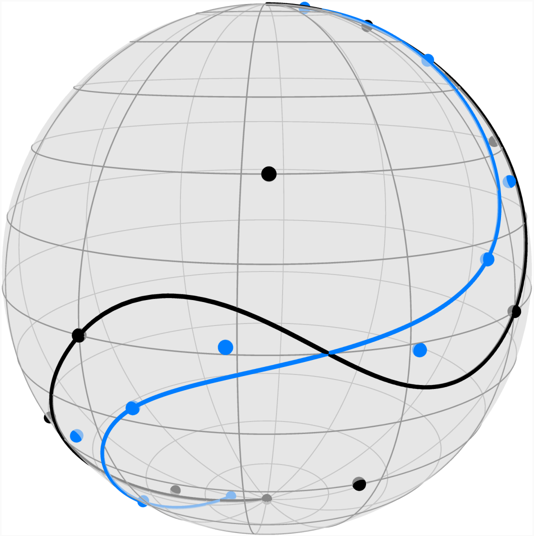
Interpolation by Bézier Curves with Minimal Acceleration.



A comp. Bezier curve (black) and its mnimizer (blue).

Approximation by Bézier Curves with Minimal Acceleration.

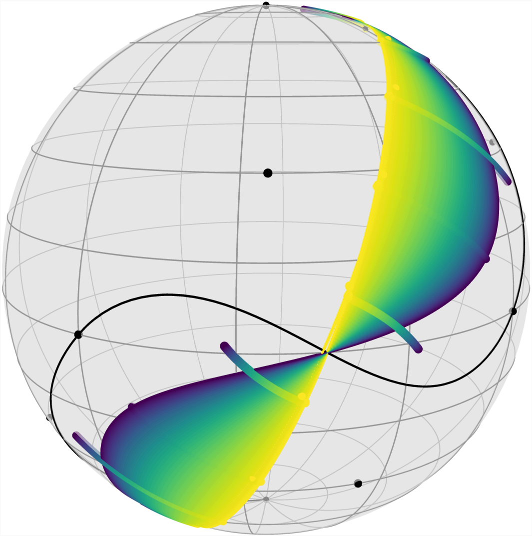
In the following video λ is slowly decreased from 10 to 0.



The initial setting, $\lambda = 10$.

Approximation by Bézier Curves with Minimal Acceleration.

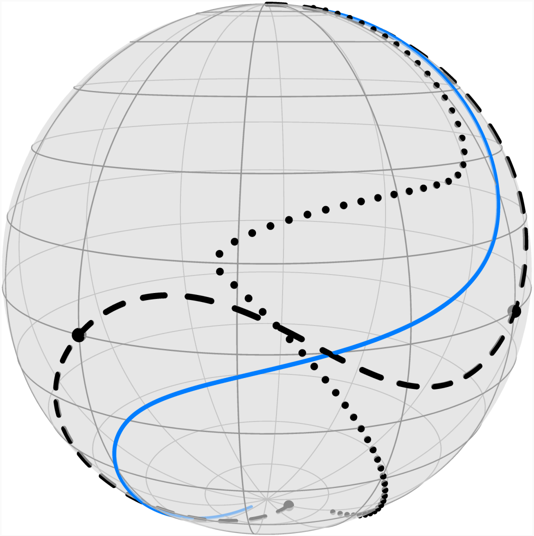
In the following video λ is slowly decreased from 10 to 0.



Summary of reducing λ from 10 (violet) to zero (yellow).

Comparison to Previous Approach

[Gousenbourger, Massart, Absil, 2018]

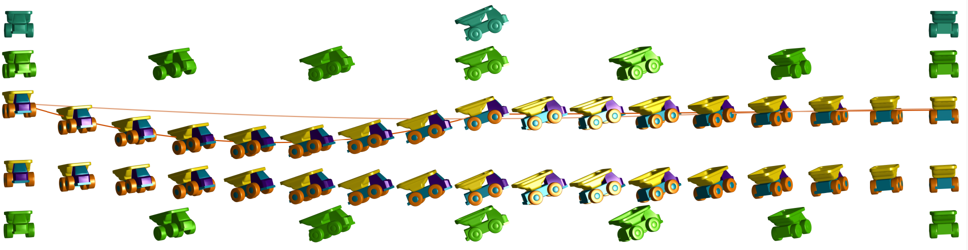


This curve (dashed) is "too global" to be solved in a tangent space (dotted) correctly, while our method (blue) still works.

An Example of Rotations $\mathcal{M} = \mathrm{SO}(3)$

Initialization with approach from composite splines

[Gousenbourger, Massart, Absil, 2018]



Our method outperforms the approach of solving linear systems in tangent spaces, but their approach can serve as an initialization.

Summary

We investigated a model to minimize the acceleration of a Bézier curve

- using second order differences
- employing Jacobi fields
- using a gradient descent w.r.t. the control points

The algorithm will be published in `Manopt.jl` a Julia Package available at `http://manoptjl.org`. Goal:

Being able to use an(y) algorithm for a(ny) model directly on a(ny) manifold easily and efficiently.

in an open source programming language.

Example

```
xOpt = steepestDescent(N, F,∇F, x0)
```

Selected References

- Arnould, A.; Gousenbourger, P.-Y.; Samir, C.; Absil, P.-A.; Canis, M. (2015). "Fitting Smooth Paths on Riemannian Manifolds : Endometrial Surface Reconstruction and Preoperative MRI-Based Navigation". *GSI2015*. Ed. by F.Nielsen; F.Barbaresco. Springer International Publishing, pp. 491–498. DOI: 10.1007/978-3-319-25040-3_53.
- Bergmann, R.; Gousenbourger, P.-Y. (2018). "A variational model for data fitting on manifolds by minimizing the acceleration of a Bézier curve". *Frontiers in Applied Mathematics and Statistics*. DOI: 10.3389/fams.2018.00059. arXiv: 1807.10090.
- Boumal, N.; Absil, P. A. (2011). "A discrete regression method on manifolds and its application to data on SO(n)". *IFAC Proceedings Volumes (IFAC-PapersOnline)*. Vol. 18. PART 1, pp. 2284–2289. DOI: 10.3182/20110828-6-IT-1002.00542.
- Gousenbourger, P.-Y.; Massart, E.; Absil, P.-A. (2018). "Data fitting on manifolds with composite Bézier-like curves and blended cubic splines". *Journal of Mathematical Imaging and Vision*. accepted. DOI: 10.1007/s10851-018-0865-2.
- Samir, C.; Absil, P.-A.; Srivastava, A.; Klassen, E. (2012). "A Gradient-Descent Method for Curve Fitting on Riemannian Manifolds". *Foundations of Computational Mathematics* 12.1, pp. 49–73. DOI: 10.1007/s10208-011-9091-7.