# Numerical Differentiation of Sampled Data

Nathan Brack

Project course in Scientific Computing
supervised by Michael Hanke

## Introduction

The task is to investigate on different numerical methods to estimate the derivative of sampled noisy data. Consider an interval $[a, b]$ with $a = x_0 < \cdots < x_n = b$. Noisy data points $y_i$ of a function $g : [a, b] \to \mathbb{R}$ are given, i. e. $y_i \approx g(x_i)$. We are looking for a good approximation $u \approx g'$.

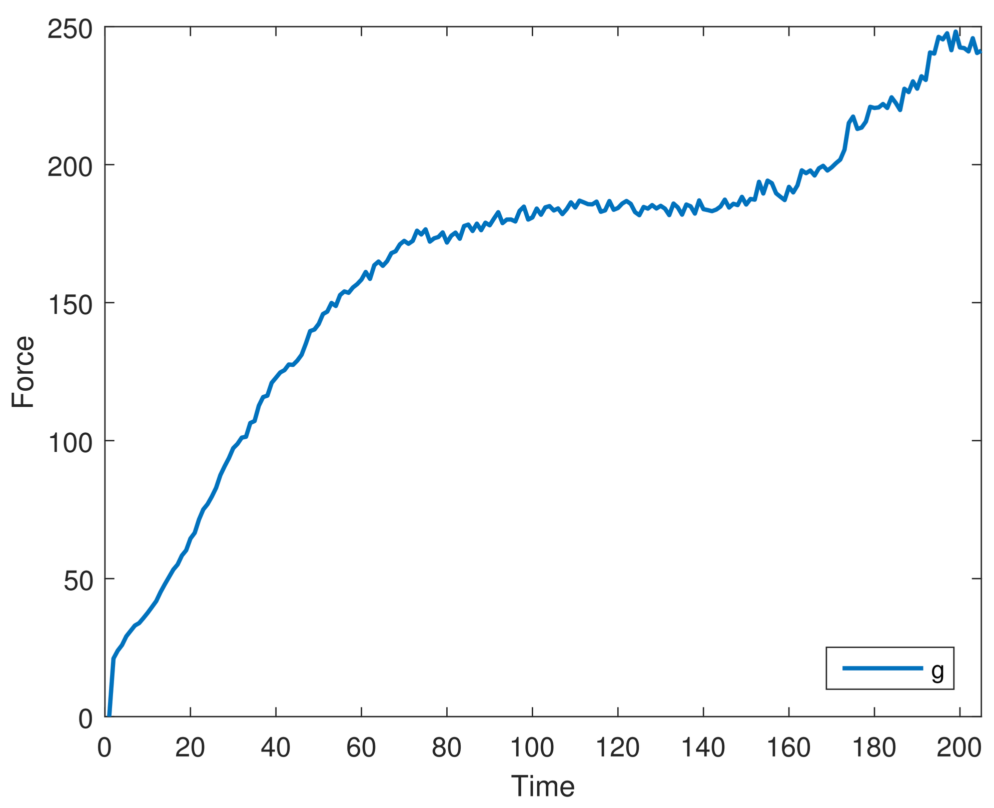


**Fig. 1:** Actual measurements of the force in a cell

## Problem

Numerical integration is numerically stable whereas its inverse problem, numerical differentiation, is an ill-posed problem. The result does not depend continously on the errors in the data. Approximating the derivative by Finite Differences, i. e. by building the differential quotient, leads to undesired oscillation.

One can regularize the derivative as follows

- Take the derivative of an appropriate interpolant (*Least Square Polynomial*)
- "Smoothen" the data first and differentiate then (*Convolution Smoothing*)
- Introduce a **Regularization Parameter** $\alpha$ to control the smoothness of the solution (*Tychonov Regularization, Smoothing Spline*)

## Results

In summary all the methods except *Improved Finite Differences* seem to perform quite well.
An obvious disadvantage of the first three methods is that the approximation is only defined on a subinterval of the form $[a + d, b - d]$ (**Fig. 3**).
The *Tychonov Regularization* and the *Smoothing Spline* deliver almost the same approximation (**Fig. 4**). This makes sense, since both control the $L^2$-norm of the derivative of the approximation $u$ (for the *Tychonov Regularization* we have choosen $k = 1$ and for the *Smoothing Spline*, note that $||f''||_2 = ||u'||_2$).

The following table shows the results for the determination of $\alpha$ rounded to three significant digits.

| $\alpha$ | Tychonov Reg. | Smoothing Sp. |
|---|---|---|
| L-Curve Criterion | 16600 | 85 |
| Cross Validation | 15700 | 125 |
| Discreprancy Pr. | 5200 | 207 |

For the *Discreprancy Principle* one needs to estimate $\delta$ respectively $||\eta||$. By **Fig. 1** we choose $\delta = 4$. Heuristically from **Fig. 2**, where the *L-Curve Criterion* is applied to the *Tychonov Regularization*, it seems reasonable to choose $||\eta|| = 2500$ ($\approx$ Error at the dot).

## Methods

- **Improved Finite Differences** Instead of the simple Finite Differences approach $u_{i+1/2} \approx \frac{y_{i+1}-y_i}{x_{i+1}-x_i}$ one can include more points into a Finite Difference formula to achieve a cancellation of errors.
- **Least Square Polynomial** To get an approximation $u(x_i)$, take $m$ points in the neighbourhood of $x_i$ and fit a polynomial of degree $p < m$ trough these points in a Least Square sense by solving the corresponding normal equations.
- **Convolution Smoothing** Take a symmetric mollifier function $\rho(x)$ supported on $(-1, 1)$ and peaked at 0 such that $\int \rho(x)dx = 1$. Let $p$ be the piecewise linear intepolant of the data points $(x_i, y_i)$. We have an approximation of $g$ by the following convolution

$$1/d \int_{a-d}^{b+d} \rho\left(\frac{x-s}{d}\right) p(s)ds$$

where $2d$ is the support of the integral. One can consider this approximation a smoothened version of $g$ on which we can now simply apply the Finite Difference method.

- **Tychonov Regularization** One can write $(Au)(x) = \int_a^x u(t)dx = g(x) - g(a)$. Discretize $Au$ and, for a given $\alpha > 0$, consider

$$E(u) = ||Au - \hat{y}||^2 + \alpha||D_k u||^2$$

where $\hat{y}_i = y_i - y_0$ and $D_k u$ stands for the discretized k-th derivative of u, where $k = 0, 1, 2$, to control the smoothness of the solution. We then minimize the functional $E(u)$ by solving its normal matrix equations.

- **Smoothing Spline** For a given $\alpha > 0$, consider

$$\Phi(f) = \frac{1}{n-1}\sum_{i=1}^{n-1}(y_i - f(x_i))^2 + \alpha||f''||_2$$

The minimizer $f_{min}$ of this functional is a natural cubic spline, i.e. piecewise polynomial of degree 3 and twice continously differentiable, satisfying

$$f'''_{min}(x_i+) - f'''_{min}(x_i-) = \frac{1}{\alpha(n-1)}(y_i - f_{min}(x_i))$$

## Regularization Parameter $\alpha$

The latter two methods require the determination of a suitable $\alpha$. The bigger we choose $\alpha$, the less our solutions fits the data points $y_i$ (increasing the Error) but the more smooth our solution gets (decreasing the Penalty Term).

- **L-Curve Criterion** Provides a good deal between keeping small the error as well as the penalty term by loglog-plotting these qunatities against each other and picking the point of highest curvature.

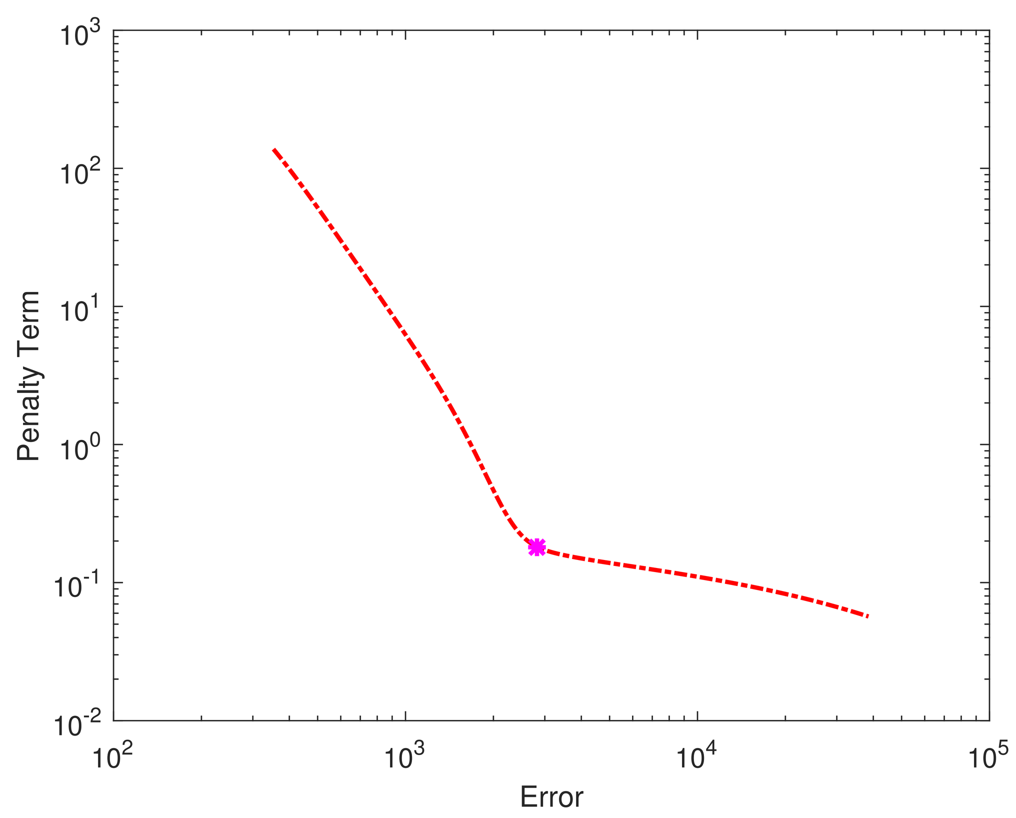


**Fig. 2:** The choosen $\alpha$ corresponds to the dot

- **Cross Validation** The statistical idea is the following: Let $X = \{(x_0, y_0), \cdots, (x_n, y_n)\}$. For $1 \leq i \leq n-1$, perform the method on the set $X \backslash \{(x_i, y_i)\}$ and validate it on $(x_i, y_i)$, that means choose the $\alpha_i$ such that the error with respect to $(x_i, y_i)$ is minimal. Finally average over all the $\alpha_i$

$$\alpha = \frac{1}{n-1}(\alpha_1 + \cdots + \alpha_{n-1})$$

- **Discrepancy Principle**
  - *Tychonov Regularization* If $\eta$ is the error vector, i.e. $\eta_i = y_i - g(x_i)$, we choose $\alpha$ such that the minimizer $u_{min}$ of the functional $E(u)$ satisfies

$$||Au_{min} - \hat{y}||^2 = ||\eta||^2$$

  - *Smoothing Spline* If $\delta$ is the known level of noise, i.e. $|y_i - g(x_i)| \leq \delta$, we choose $\alpha$ such that the minimizer $f_{min}$ of the functional $\Phi(f)$ satisfies

$$\frac{1}{n-1}\sum_{i=1}^{n-1}(y_i - f_{min}(x_i))^2 = \delta^2$$

## Figures

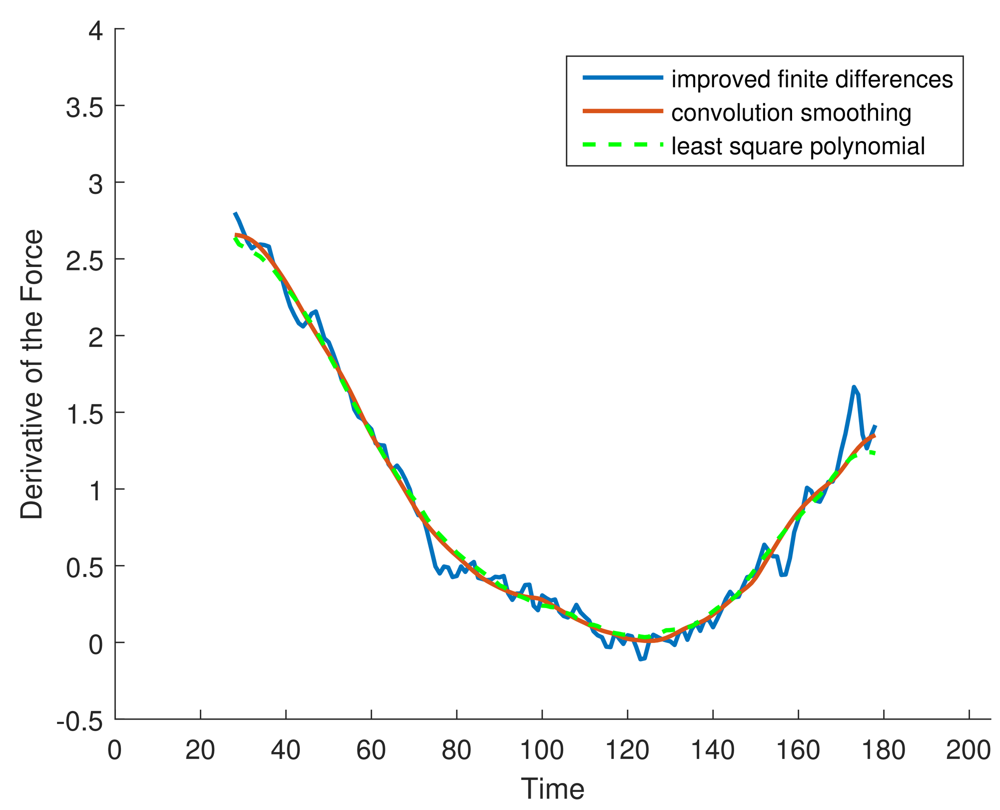


**Fig. 3:** For the *Least Square Polynomial* method the degree of the polynomial is choosen to be $p = 2$.
For the *Convolution Smoothing* the mollifier is choosen to be $\rho(x) = \chi_{(-1,1)} ce^{\frac{1}{x^2-1}}$.

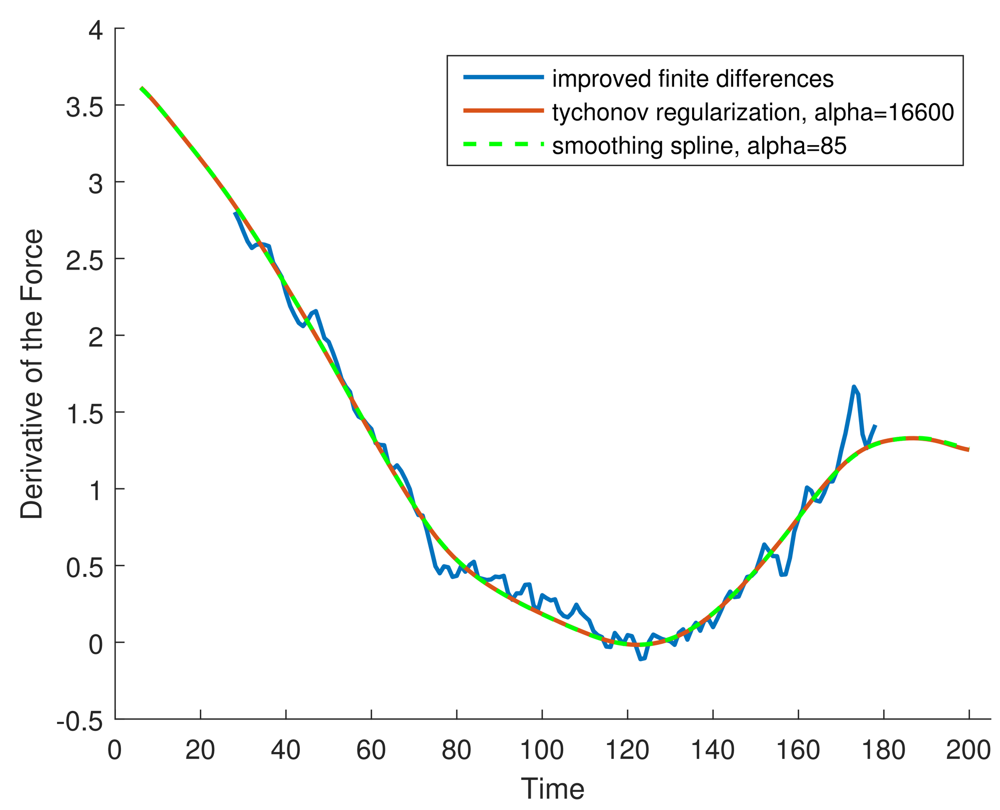


**Fig. 4:** For both methods the $\alpha$ was determined by the *L-Curve Criterion*.
For the *Tychonov Regularization* the order of the derivative is choosen to be $k = 1$.

## References

[1] Ian Knowles, Robert J. Renka: *Methods for Numerical Differentiation of Noisy Data*

[2] Martin Hanke, Otmar Scherzer: *Inverse Problems Light: Numerical Differentiation*

[3] Heinz Werner Engl, Martin Hanke, A. Neubauer: *Regularization of Inverse Problems*