## Theoretical background for cubic B-spline evaluation

Sven Proppert

June 23, 2018

## 1 Definition of B-spline basis functions

Consider a given dataset of m points  $y_0(x_0)$ ,  $y_1(x_1)$ , ...,  $y_m(x_m)$  on an interval  $[L_1 : L_2]$  (black crosses in fig.1). In order to perform the interpolation, first that interval is split into segments represented by n equidistant knots  $t_i$  (i = 0, 1, ..., n - 1) marked as black circles. As implemented in the GNU Scientific Library for B-spline-interpolation, the n+2 basis functions  $s_i(x)$  of order k are segmentwise recursively defined following equations 1 and 2. For cubic B-splines, k goes from 1 to 4.

$$s_{i,1}(x) = \begin{cases} 1 & : t_i \le x < t_{i+1} \\ 0 & : else \end{cases}$$
(1)

$$s_{i,k}(x) = \frac{x - t_i}{t_{i+k-1} - t_i} s_{i,k-1}(x) + \frac{t_{i+k} - x}{t_{i+k} - t_{i+1}} s_{i+1,k-1}(x)$$
(2)

## 2 Interpolation and least squares fitting (LSF) as implemented in rapidSTORM

The interpolation to the dataset is computed as in eq.3 and represents a set of linear equations with  $\vec{y} = (y_0(x_0), y_1(x_1), \ldots, y_m(x_m))$  and  $\vec{c} = (c_0, c_1, \ldots, c_n, c_{n+1}, c_{n+2})$ . Matrix notation (eq.4) indicates, that S is of diagonal form.

$$y_m(x_m) = \sum_i s_{i,k}(x_m)c_i$$
$$\vec{y} = S \cdot \vec{c}$$
(3)



Figure 1: This diagram (arbitrary units) aims to visualise exemplarily how to handle data with cubic B-spline interpolation. The black crosses represent seventeen given data points while the rainbow-colored pseudo-Gaussians at the bottom of the diagram are plots of the B-spline basis functions  $s_{i,4}(x)$  (i = 0, 1, ..., 11) with the n = 9 knots indicated by black circles. Evaluation of eq.5 leads to amplitude factors for each basis function. The superposition with given amplitudes results in the interpolation represented by the red parabola.



In equation 3 the vector  $\vec{c}$  is unknown and is object to least squares fitting, which is performed following eq.5.

$$\vec{c} = (S^T \cdot S)^{-1} \cdot S^T \cdot \vec{y} \tag{5}$$

## 2.1 Derivation of the formula for LSF (eq.5)

The fit residuals are defined as:  $\vec{r}(\vec{c}) = \vec{y} - S \cdot \vec{c}$  with the *i*-th component of r being:  $r_i(c_j) = y_i - \sum_j S_{ij} \cdot c_j$ . The least squares fitting criteria is fulfilled when

$$\left|\vec{r}(\vec{c})^2\right| = \min \Leftrightarrow \frac{\partial \vec{r}(\vec{c})^2}{\partial \vec{c}} = 0$$
 (6)

Evaluation of the right side of eq.6 for  $r_i$  gives

$$0 = \sum_{i} \frac{\partial \left( r_i(c_j)^2 \right)}{\partial c_j} = \sum_{i} \frac{\partial r_i(c_j)}{\partial c_j} \cdot 2r_i(c_j)$$
$$= \sum_{i} -S_{ij} \cdot 2r_i(c_j) = \sum_{i} S_{ij} \cdot \left[ y_i - \sum_k S_{ik} \cdot c_k \right]$$
$$= \sum_{i} S_{ij} \cdot y_i - \sum_{i} S_{ij} \cdot \sum_k S_{ik} \cdot c_k$$

$$\sum_{i} S_{ij} \cdot y_{j} = \sum_{k} \left[ \sum_{i} S_{ij} \cdot S_{ik} \right] \cdot c_{k}$$
$$\sum_{i} S_{ji}^{T} \cdot y_{i} = \sum_{k} \left[ \sum_{i} S_{ji}^{T} \cdot S_{ik} \right] \cdot c_{k}$$
$$= \sum_{k} \left( S^{T} \cdot S \right)_{jk} \cdot c_{k}$$
$$\left( S^{T} \cdot \vec{y} \right)_{j} = \left[ \left( S^{T} \cdot S \right) \vec{c} \right]_{j}$$
(7)

which, rewritten in vector notation leads to the form of eq.5:

$$S^{T}\vec{y} = (S^{T} \cdot S)\vec{c}$$
  

$$\Leftrightarrow (S^{T} \cdot S)^{-1}S^{T}\vec{y} = \mathbb{1}\vec{c} \quad . \tag{8}$$