

# ISOGOMETRIC ANALYSIS

THEORY AND PRACTICE

Ahmed Ratnani

Copyright © 2021 A. Ratnani

PUBLISHED BY UM6P UNIVERSITY

UM6P.MA

Licensed under the Creative Commons Attribution-NonCommercial 3.0 Unported License (the “License”). You may not use this file except in compliance with the License. You may obtain a copy of the License at <http://creativecommons.org/licenses/by-nc/3.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

*First printing, October 2021*

## Foreword

This book was created and is still evolving out of my lectures at the graduate level, that I taught at TUM and still teaching at UM6P (Master TIUF, Master MSD).

Last but not least, I would like to thank my graduate students and colleagues, for their useful comments and discussions.

Dr. Ahmed Ratnani

### TODO:

- Historical notes
- Exercises
- Applications
- Linear Solvers
- Add the lecture Approximation and Learning theory?

# Contents

<b>I</b>	<b>Computer Aided Design</b>	
<b>1</b>	<b>Introduction</b> .....	<b>11</b>
1.1	Implicit form	11
1.2	Parametric curves	11
1.3	Power basis form of a curve	13
1.4	Problems	13
<b>2</b>	<b>Bézier curves</b> .....	<b>15</b>
2.1	Bernstein polynomials	15
2.2	Bézier curves	16
2.3	DeCasteljau Algorithm	21
2.4	Conversion from/to monomial form	22
2.5	Rational Bézier curves	23
2.6	Composite Bézier curves	23
2.7	Problems	25

<b>3</b>	<b>B-Splines</b> .....	<b>27</b>
3.1	Knot vector families	28
3.2	Examples	29
3.3	B-Splines properties	38
3.4	Derivatives of B-Splines	40
3.5	Problems	41
<b>4</b>	<b>Cardinal B-Splines</b> .....	<b>42</b>
4.1	Cardinal B-Splines	42
4.1.1	Cardinal B-Splines properties .....	44
4.2	Cardinal B-Spline series	46
4.2.1	Scaled and translated Cardinal B-Splines .....	46
4.2.2	Cardinal Splines .....	46
4.3	Problems	47
<b>5</b>	<b>B-Splines curves</b> .....	<b>48</b>
5.1	B-Splines curves	48
5.2	Derivative of a B-spline curve	52
5.3	Rational B-Splines ( <i>NURBS</i> ) curves	53
5.3.1	Modeling conics using <i>NURBS</i> .....	53
5.4	Fundamental geometric operations	59
5.4.1	Knot insertion .....	59
5.4.2	Degree elevation .....	62
5.5	Problems	64
<b>6</b>	<b>Historical Notes</b> .....	<b>66</b>

## II

## Approximation theory for B-Splines

<b>7</b>	<b>Divided differences</b> .....	<b>69</b>
7.1	Lagrange interpolation	69
7.1.1	Newton form and Neville's algorithm .....	70
7.2	Hermite interpolation	71
7.3	Divided differences	71
7.4	Problems	73
<b>8</b>	<b>Schoenberg space of Spline functions</b> .....	<b>74</b>
8.1	Basic Splines	74
8.1.1	Smoothness of a B-Spline .....	75
8.2	Spline functions	75
8.3	Dual functionals	77
8.4	Problems	78

<b>9</b>	<b>Spline Approximation</b>	<b>79</b>
<b>9.1</b>	<b>Introduction</b>	<b>79</b>
<b>9.2</b>	<b>Examples</b>	<b>79</b>
9.2.1	Piecewise linear interpolation	79
9.2.2	Variation diminishing spline approximation	80
<b>9.3</b>	<b>Quasi-Interpolation</b>	<b>83</b>
9.3.1	General recipe for quasi-interpolants	83
9.3.2	Examples	83
9.3.3	Computing the coefficients functionals	85
<b>9.4</b>	<b>Global approximation</b>	<b>86</b>
9.4.1	Interpolation	86
9.4.2	Histopolation	87
9.4.3	Least-square approximation	88
<b>9.5</b>	<b>Approximation with Quasi-Interpolation</b>	<b>90</b>
9.5.1	Quasi-interpolation and reproduction of polynomials	91
9.5.2	Examples	92
<b>9.6</b>	<b>Approximation power of Splines</b>	<b>92</b>
<b>9.7</b>	<b>Problems</b>	<b>92</b>
<b>10</b>	<b>Historical Notes</b>	<b>93</b>

### III

## Finite Elements method

<b>11</b>	<b>Functional Analysis</b>	<b>96</b>
<b>11.1</b>	<b>Notations and Preliminaries</b>	<b>96</b>
<b>11.2</b>	<b>Sobolev spaces</b>	<b>97</b>
11.2.1	The Sobolev space $W^{s,m}$	97
11.2.2	The Sobolev space $H^m$	98
11.2.3	Inequalities	98
<b>11.3</b>	<b>The Sobolev space <math>H(\text{curl}, \Omega)</math></b>	<b>98</b>
<b>11.4</b>	<b>The Sobolev space <math>H(\text{div}, \Omega)</math></b>	<b>98</b>
<b>11.5</b>	<b>DeRham sequences</b>	<b>98</b>
11.5.1	Exact discrete DeRham sequence	99
11.5.2	Space decompositions	99
<b>11.6</b>	<b>Problems</b>	<b>100</b>
<b>12</b>	<b>Galerkin methods</b>	<b>101</b>
<b>12.1</b>	<b>Abstract framework</b>	<b>101</b>
<b>12.2</b>	<b>Galerkin Approximation</b>	<b>102</b>
12.2.1	Convergence under coercivity	102
12.2.2	Convergence under <i>inf-sup</i> conditions	103
12.2.3	The three basic aspects of the Finite Elements method	104
12.2.4	Examples	105

<b>12.3</b>	<b>Saddle-point problems</b>	<b>107</b>
12.3.1	Examples	108
12.3.2	Galerkin approximation	110
12.3.3	Examples	112
<b>12.4</b>	<b>Problems</b>	<b>114</b>
<b>13</b>	<b>Historical Notes</b>	<b>115</b>

## IV

## Isogeometric Analysis

<b>14</b>	<b>Isogeometric Finite Elements</b>	<b>118</b>
<b>14.1</b>	<b>Sobolev estimations under h-refinement</b>	<b>118</b>
14.1.1	Approximation properties without mapping	118
14.1.2	Approximation properties with mapping	121
<b>14.2</b>	<b>Galerkin approximation</b>	<b>122</b>
14.2.1	Case without mapping	122
14.2.2	Case with mapping	124
<b>14.3</b>	<b>Exact DeRham sequences</b>	<b>127</b>
14.3.1	Commuting diagrams	130
14.3.2	Examples without mapping	133
14.3.3	Pullbacks	137
14.3.4	Examples with mapping	138
<b>14.4</b>	<b>Problems</b>	<b>138</b>
<b>15</b>	<b>Historical Notes</b>	<b>139</b>

## V

## Linear Algebra

<b>16</b>	<b>Kronecker Algebra</b>	<b>142</b>
<b>16.1</b>	<b>Kronecker algebra</b>	<b>142</b>
16.1.1	Kronecker sum	145
16.1.2	Solving $AX + XB = C$	145
16.1.3	Solving $AXB = C$	146
16.1.4	Solving $\sum_{i=1}^r A_i X B_i = C$	146
<b>16.2</b>	<b>Problems</b>	<b>146</b>
<b>17</b>	<b>Historical Notes</b>	<b>147</b>

## VI

## Applications

<b>18</b>	<b>Navier Stokes</b>	<b>150</b>
<b>18.1</b>	<b>Problems</b>	<b>150</b>

<b>19</b>	<b>Historical Notes</b> .....	<b>151</b>
	<b>Bibliography</b> .....	<b>152</b>
	<b>Articles</b>	<b>152</b>
	<b>Books</b>	<b>153</b>





# Computer Aided Design

<b>1</b>	<b>Introduction .....</b>	<b>11</b>
1.1	Implicit form	
1.2	Parametric curves	
1.3	Power basis form of a curve	
1.4	Problems	
<b>2</b>	<b>Bézier curves .....</b>	<b>15</b>
2.1	Bernstein polynomials	
2.2	Bézier curves	
2.3	DeCasteljau Algorithm	
2.4	Conversion from/to monomial form	
2.5	Rational Bézier curves	
2.6	Composite Bézier curves	
2.7	Problems	
<b>3</b>	<b>B-Splines .....</b>	<b>27</b>
3.1	Knot vector families	
3.2	Examples	
3.3	B-Splines properties	
3.4	Derivatives of B-Splines	
3.5	Problems	
<b>4</b>	<b>Cardinal B-Splines .....</b>	<b>42</b>
4.1	Cardinal B-Splines	
4.2	Cardinal B-Spline series	
4.3	Problems	
<b>5</b>	<b>B-Splines curves .....</b>	<b>48</b>
5.1	B-Splines curves	
5.2	Derivative of a B-spline curve	
5.3	Rational B-Splines ( <i>NURBS</i> ) curves	
5.4	Fundamental geometric operations	
5.5	Problems	
<b>6</b>	<b>Historical Notes .....</b>	<b>66</b>



# 1. Introduction

## 1.1 Implicit form

Among the different representations to describe a quarter of circle (of radius 1 and centered at the origin), one can use an implicit form using the equation

$$f(x,y) := x^2 + y^2 - 1 = 0, \quad x \geq 0 \text{ and } y \geq 0$$

In practice, implicit forms are well adapted for unbounded geometries. On the other hand, given a point, it is easy to determine if it is on the curve or a surface. Although, this is an important question, in general, we are interested in geometric manipulations in CAD systems, such as improving the resolution, smoothness and having local control of a given curve or surface. These properties are very difficult to ensure when using an implicit form; at the end, a computer needs discrete values in order to plot or manipulate a CAD object, which is in opposition with the analytical form represented by an equation.

## 1.2 Parametric curves

Another way of representing a quarter of circle is to consider the parametric curve define in 1.1, where  $\theta \in [0, \frac{\pi}{2}]$

$$\mathcal{C}(\theta) = \begin{bmatrix} x(\theta) \\ y(\theta) \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad (1.1)$$

in which case we can compute its derivative as

$$\mathcal{C}'(\theta) = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix} \quad (1.2)$$

The tangent at  $\theta = 0$  and  $\theta = \frac{\pi}{2}$  are then given by  $\mathcal{C}'(\theta := 0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  and  $\mathcal{C}'(\theta := \frac{\pi}{2}) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ .

Another way to represent the quarter of circle, is to introduce the variable  $t := \tan \frac{\theta}{2}$ , then we get

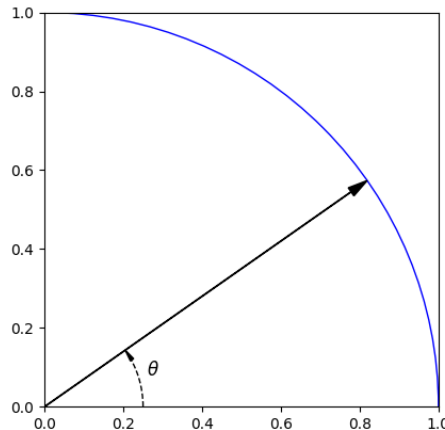


Figure 1.1: Quarter of circle of radius 1, using polar coordinates.

the following rational form 1.3, where  $t \in [0, 1]$ ,

$$\mathcal{C}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \frac{1-t^2}{1+t^2} \\ \frac{2t}{1+t^2} \end{bmatrix} \quad (1.3)$$

There are different advantages for having a parametric form:

- Extending a 2D curve to 3D,
- Ideal for bounded curves and surfaces,
- Natural orientation of curves and surfaces,
- Appropriate representation for design on a computer. In some particular cases, the coefficients in a parametric form possess interesting geometric significance,
- Computing a point on a curve is easy, while finding its parametric value is in general a non-linear problem,

**R** Notice that in some cases, this representation may introduce singularities that are unrelated to the described geometry.

**R** Computing the derivatives (tangent or normal vectors at the extremities) needs an evaluation. We shall be more interested in representation where such informations can be extracted directly from the parametric representations.

Although one can use any parametric representation for CAD systems, it is important to restrict to a representation that is

- capable of reproducing a *wild class* of curves/surfaces
- easy to implement
- efficient and accurate evaluation
- numerical stable evaluation
- *easy* evaluation of points and their derivatives
- evaluation complexity of the same or comparable order as polynomials
- small memory storage

### 1.3 Power basis form of a curve

If we restrict our representation to polynomials, one naive way would be to use power basis, *i.e.* monomials as in 1.4

$$\mathcal{C}(t) = \sum_{i=0}^n \mathbf{a}_i t^i \quad (1.4)$$

where  $\mathbf{a}_i = \begin{bmatrix} a_i^x \\ a_i^y \\ a_i^z \end{bmatrix}$  are 2D or 3D vectors.

Using the Taylor expansion, we get  $\mathcal{C}'(t)|_{t=0} = i! \mathbf{a}_i$ , which leads to  $\mathbf{a}_i = \frac{\mathcal{C}'(t)|_{t=0}}{i!}$ . We then have a direct access to the derivative at  $t = 0$ , but only at this point.

The evaluation of such a form can be done using the Horner algorithm as described in 1.5:

$$\mathcal{C}(t) = a_0 + t(a_1 + t(a_2 + t(\cdots + t(a_{n-1} + ta_n)))) \quad (1.5)$$

---

**Algorithm 1:** HORNER: Evaluation of a polynomial defined by its coefficients  $\mathbf{a}$  at  $x$ .

---

**Input** :  $\mathbf{a}, x$

**Output** :  $v$

```

1  $n \leftarrow \text{len}(\mathbf{P}) - 1$ 
2  $v \leftarrow a[n]$ 
3 for  $i \leftarrow n - 1$  to 0 do
4   |  $v \leftarrow vx + a[i]$ 
5 end
6 return  $v$ 

```

---

**R** The Horner algorithm is numerically unstable for high degrees, more details can be found in [7] [8] [12] [2].

**R** The use of power basis form is not suited for shape and geometric design; the coefficients do not have any geometric meaning and modifying them do not allow for a good control of a curve or a surface,

### 1.4 Problems

**Exercise 1.1** Consider the two parametric representations of the circular arc given by Eqs. (1.1) and (1.3). Using Eq. (1.1), compute the curve point at  $u = \frac{\pi}{4}$  and, using Eq. (1.3), the point at  $t = \frac{1}{2}$ . Explain the results. ■

**Exercise 1.2** Compute the acceleration vector,  $\mathcal{C}''(u)$ , for Eq. (1.1). Explain the result. ■

**Exercise 1.3** Consider the parabolic arc  $\mathcal{C}''(u) = (x(u), y(u)) = (-1 - u + 2u^2, -2u, u^2)$  for  $0 \leq u \leq 1$ . Plot the curve  $\mathcal{C}$ , then apply the transformations

- a rotation of  $\frac{\pi}{2}$  about the origin. We recall that the rotation matrix (applied from the left)

is

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

- a translation with the vector  $(-1, -1)$ .

Compute the implicit equation of the associated parabola. ■

**Exercise 1.4** Determine formulas for the number of additions and multiplications necessary to compute a point on an  $n$ -th-degree three-dimensional power basis curve. ■

**Exercise 1.5** Construct a cubic power basis curve with a loop. Hint: think about what end-points and end derivatives,  $\mathcal{C}'(0)$  and  $\mathcal{C}'(1)$ , are necessary. ■

**Exercise 1.6** Construct a cubic power basis curve with a cusp. Hint: think about  $\mathcal{C}'(u)$  and  $\mathcal{C}''(u)$ . Sketch what  $x'(u)$ ,  $y'(u)$ ,  $x''(u)$ , and  $y''(u)$  need to look like as functions of  $u$ . Determine a suitable  $\mathcal{C}''(u)$ , and then integrate to obtain  $\mathcal{C}'(u)$  and  $\mathcal{C}(u)$ . ■

**Exercise 1.7** Construct a cubic power basis curve with an inflection point. ■

**Exercise 1.8** Compare the CPU cost for an evaluation for the presented representations. ■

**Exercise 1.9** Compute the CPU cost for the Horner algorithm. ■

**Exercise 1.10** Write a Python implementation for the Horner algorithm. ■

**Exercise 1.11 — Parallel evaluation.** Let us consider a polynomial  $P$  of degree  $n$ , defined as  $P(t) = \sum_{i=0}^n a_i t^i$ . Using the form

$$P(t) = \sum_{j=0}^{k-1} x^j P_j(x^k)$$

as a splitting into  $k$  independent parts, write a Python code for the parallel evaluation and compute its complexity. ■

## 2. Bézier curves

### 2.1 Bernstein polynomials

In 1885, Weierstrass proved that the set of polynomial functions on  $[0, 1]$  is dense in  $\mathcal{C}([0, 1])$ . In 1912, Sergei Bernstein gave a simple proof to this result by introducing the now-famous Bernstein polynomials 2.1:

$$B_k^n(x) = \binom{n}{k} x^k (1-x)^{n-k}, \quad \text{where } k \in [0, n] \text{ and } x \in [0, 1] \quad (2.1)$$

In figure Fig. 2.1, we plot Bernstein polynomials of degrees 1 to 5. From these plots, we see that Bernstein polynomials are positive, the first and last polynomials are equal to 1 on the 0 and 1 respectively. More properties will be discussed in the sequel.

#### Properties of Bernstein polynomials

- $B_k^n(x) \geq 0$ , for all  $k \in [0, n]$  and  $x \in [0, 1]$  [positivity]
- $\sum_{k=0}^n B_k^n(x) = 1$ , for all  $x \in [0, 1]$  [partition of unity]
- $B_0^n(0) = B_n^n(1) = 1$
- $B_k^n$  has exactly one maximum on the interval  $[0, 1]$ , at  $\frac{k}{n}$
- $(B_k^n(x))_{0 \leq k \leq n}$  is symmetric with respect to  $x = \frac{1}{2}$  [symmetry]
- Bernstein polynomials can be defined recursively using the formulae 2.2

$$B_k^n(x) = (1-x)B_k^{n-1}(x) + xB_{k-1}^{n-1}(x) \quad (2.2)$$

where we assume  $B_k^n(x) = 0$  is  $k < 0$  or  $k > n$

- Bernstein derivatives can be computed using the formulae 2.3

$$B_k^{n'}(x) = n(B_{k-1}^{n-1}(x) - B_k^{n-1}(x)) \quad (2.3)$$

using the same assumption as before.

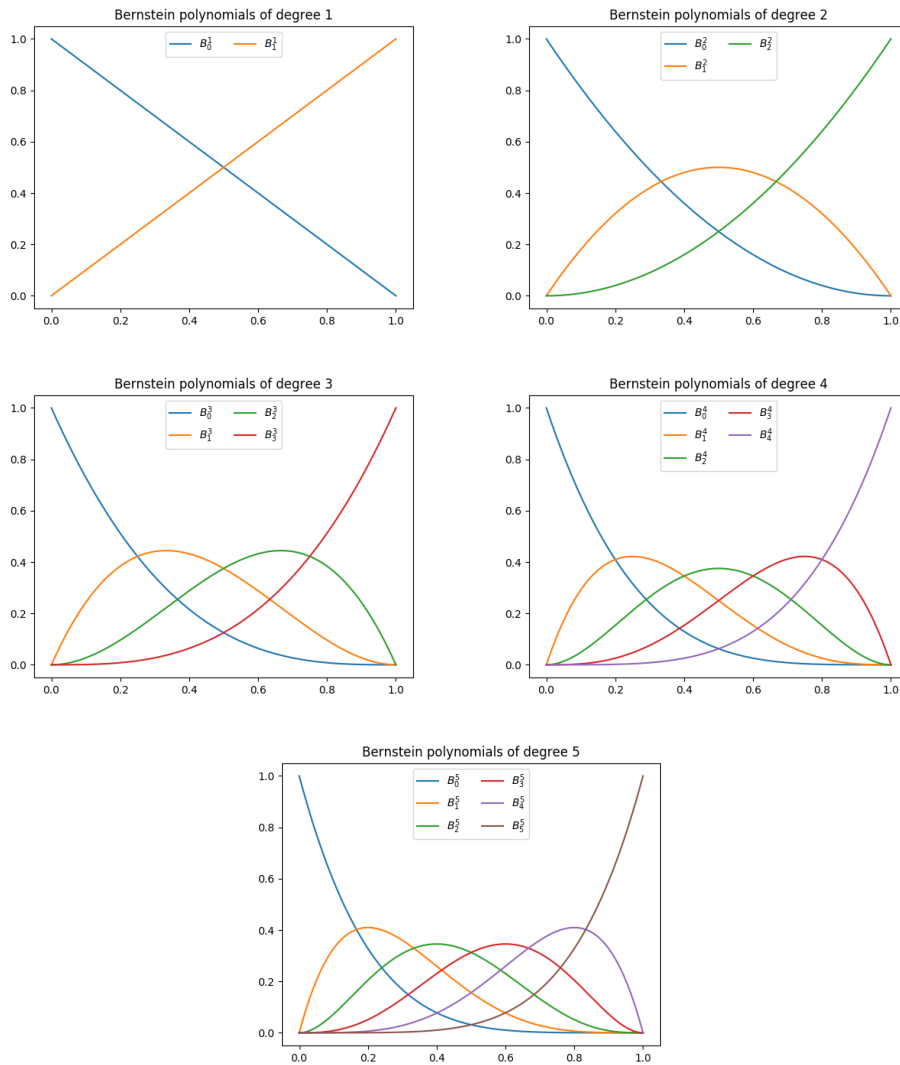


Figure 2.1: Bernstein polynomials of degree  $n = 1, 2, 3, 4, 5$ .

## 2.2 Bézier curves

Rather than taking the power basis as in 1.4, Pierre Bézier used Bernstein polynomials, in the 1960s while working at *Renault*. This leads to the following definition of a polynomial curve 2.11

$$\mathcal{C}(t) = \sum_{k=0}^n \mathbf{P}_k B_k^n(t) \quad (2.4)$$

The vector coefficients  $(\mathbf{P})_{0 \leq k \leq n}$  are called *Bézier points* or **control points**. The reason for this will become clear in the sequel.

### Examples

**Example 1.** A linear ( $n = 1$ ) Bézier curve (fig. 2.3) is defined as

$$\mathcal{C}(t) = \mathbf{P}_0 B_0^1(t) + \mathbf{P}_1 B_1^1(t)$$

which describes a straight line from  $\mathbf{P}_0$  to  $\mathbf{P}_1$ , since  $B_0^1(t) = 1 - t$  and  $B_1^1(t) = t$ .



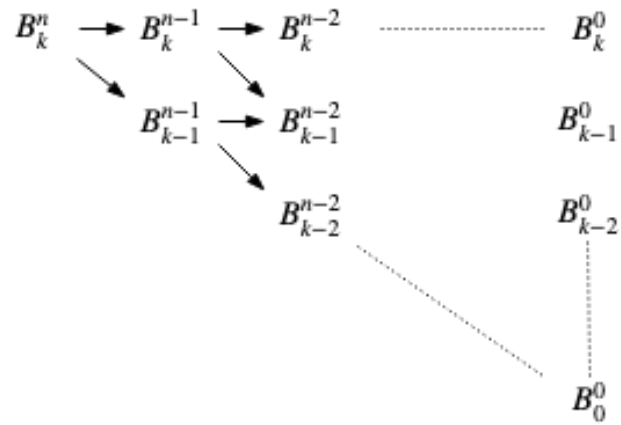


Figure 2.2: General evaluation triangular diagram for a Bernstein polynomials.

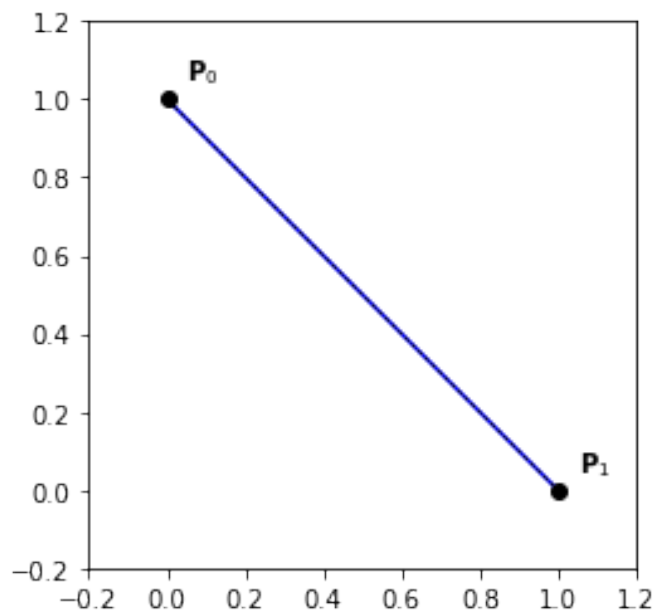


Figure 2.3: Example of a linear Bézier curve.

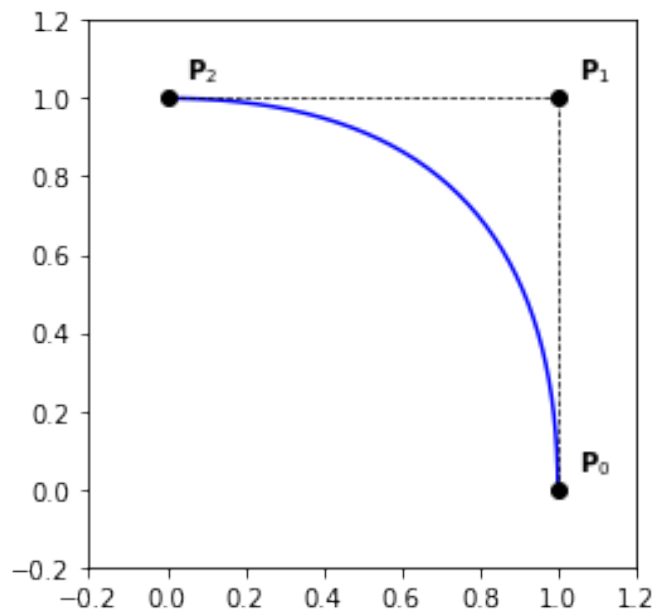


Figure 2.4: Example of a quadratic Bézier curve.

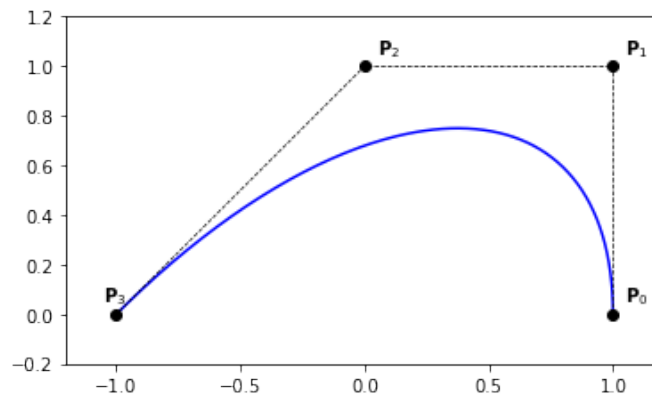


Figure 2.5: Example of a cubic Bézier curve.

**Example 2.** A quadratic ( $n = 2$ ) Bézier curve (fig. 2.4) is defined as

$$\mathcal{C}(t) = (1-t)^2\mathbf{P}_0 + 2t(1-t)\mathbf{P}_1 + t^2\mathbf{P}_2$$

We remark that

- the curve starts at  $\mathbf{P}_0$  and ends at  $\mathbf{P}_2$ ,
- the curve does not pass through the point  $\mathbf{P}_1$ ,
- the tangent directions to the curves at its extremities are parallel to  $\mathbf{P}_1 - \mathbf{P}_0$  and  $\mathbf{P}_2 - \mathbf{P}_1$ ,
- the curve is contained in the polygon (triangle) formed by  $\mathbf{P}_0\mathbf{P}_1\mathbf{P}_2$ . This polygon is called the *control polygon* and it approximates the shape of the curve.

**Example 3.** A cubic ( $n = 3$ ) Bézier curve (figures 2.5,2.6,2.7) is defined as

$$\mathcal{C}(t) = (1-t)^3\mathbf{P}_0 + 3t(1-t)^2\mathbf{P}_1 + 3t^2(1-t)\mathbf{P}_2 + t^3\mathbf{P}_3$$

We remark that

- the curves start at  $\mathbf{P}_0$  and end at  $\mathbf{P}_3$ ,

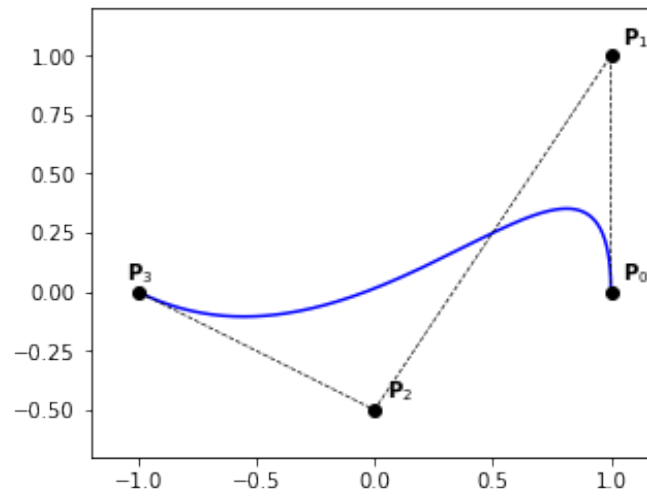


Figure 2.6: Example of a cubic Bézier curve.

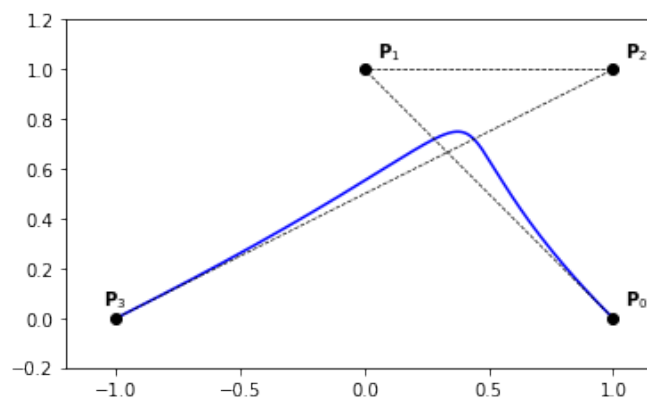


Figure 2.7: Example of a cubic Bézier curve.

- the curves do not pass through the points  $\mathbf{P}_1$  and  $\mathbf{P}_2$ ,
- the tangents directions to the curves at their extremities are parallel to  $\mathbf{P}_1 - \mathbf{P}_0$  and  $\mathbf{P}_3 - \mathbf{P}_2$ ,
- the control polygons approximate the shape of the curves,
- the curves follow the orientation of the control polygons,
- the curves are contained in the convex hulls of their control polygons (*convex hull property*).

### Derivatives of a Bézier curve

Using the formulae 2.3, we have

$$\mathcal{C}'(t) = \sum_{k=0}^n \mathbf{P}_k B_k^{n'}(t) = \sum_{k=0}^n n \mathbf{P}_k (B_{k-1}^{n-1}(x) - B_k^{n-1}(x))$$

by reordering the indices, we get 2.5

$$\mathcal{C}'(t) = \sum_{k=0}^{n-1} n (\mathbf{P}_{k+1} - \mathbf{P}_k) B_k^{n-1}(t) \quad (2.5)$$

Therefor, we get the direct acces to the first order derivatives at the extremities of a Bézier curve using the formulae 2.6

$$\begin{cases} \mathcal{C}'(0) = n (\mathbf{P}_1 - \mathbf{P}_0) \\ \mathcal{C}'(1) = n (\mathbf{P}_n - \mathbf{P}_{n-1}) \end{cases} \quad (2.6)$$

Second derivatives can also be computed directly from the control points using the formulae 2.7

$$\begin{cases} \mathcal{C}''(0) = n(n-1) (\mathbf{P}_0 - 2\mathbf{P}_1 + \mathbf{P}_2) \\ \mathcal{C}''(1) = n(n-1) (\mathbf{P}_n - 2\mathbf{P}_{n-1} + \mathbf{P}_{n-2}) \end{cases} \quad (2.7)$$

**Definition 2.2.1 —  $r^{th}$  forward difference.** Let us consider a set of (control) points  $\mathbf{P}$ . The  $r^{th}$  forward difference of  $\mathbf{P}$  is defined as

$$\Delta^r \mathbf{P}_i := \Delta^{r-1} \mathbf{P}_{i+1} - \Delta^{r-1} \mathbf{P}_i$$

with

$$\Delta \mathbf{P}_i = \Delta^1 \mathbf{P}_i := \mathbf{P}_{i+1} - \mathbf{P}_i$$

**Proposition 2.2.1 — High order derivatives.**

$$\mathcal{C}^{(r)}(t) = \frac{n!}{(n-r)!} \sum_{i=0}^{n-r} \Delta^r \mathbf{P}_i B_i^{n-r}(t) \quad (2.8)$$

Eqs (2.6) and (2.7) can be generalaized for high order derivatives. We have in fact the following result:

- R** The derivatives of a Bézier curve at its extremities up to order  $r$  depend only on the first (or last)  $r+1$  control points, and vice versa.

**Proposition 2.2.2**

$$\Delta^r \mathbf{P}_0 = \sum_{i=0}^r (-1)^{r-i} \binom{r}{i} \mathbf{P}_i$$

### Integration of Bézier curves

**Proposition 2.2.3** A primitive of A Bézier curve  $\mathcal{C}(t) = \sum_{k=0}^n \mathbf{P}_k B_k^n(t)$  has the Bézier representation

$$\left( \int \mathcal{C} \right) (t) = \sum_{k=0}^{n+1} \mathbf{Q}_k B_k^{n+1}(t) \quad (2.9)$$

where

$$\mathbf{Q}_k := \mathbf{Q}_{k-1} + \frac{1}{n+1} \mathbf{P}_{k-1} = \mathbf{Q}_0 + \frac{1}{n+1} \left( \sum_{i=0}^{k-1} \mathbf{P}_i \right)$$

Here,  $\mathbf{Q}_0$  denotes an arbitrary integration constant.

**R** Using Eq. (2.9), we have  $\int_0^1 \mathcal{C}(t) dt = \frac{1}{n+1} \left( \sum_{i=0}^n \mathbf{P}_i \right)$ .

### Properties of Bézier curves

Because of the *symmetry* of the Bernstein polynomials, we have

- $\sum_{k=0}^n \mathbf{P}_k B_k^n = \sum_{k=0}^n \mathbf{P}_k B_{n-k}^n$

Thanks to the interpolation property, for  $B_0^n$  and  $B_n^n$ , of the Bernstein polynomials at the endpoints, we get

- $\mathcal{C}(0) = \mathbf{P}_0$  and  $\mathcal{C}(1) = \mathbf{P}_n$

Using the partition unity property of the Bernstein polynomials, we get

- any point  $\mathcal{C}(t)$  is an *affine combination* of the control points

Therefore,

- Bézier curves are *affinely invariant*; *i.e.* the image curve  $\Phi(\sum_{k=0}^n \mathbf{P}_k B_k^n)$  of a Bézier curve, by an affine mapping  $\Phi$ , is the Bézier curve having  $(\Phi(\mathbf{P}_i))_{0 \leq i \leq n}$  as control points.

due to the partition unity of the Bernstein polynomials and their non-negativity,

- any point  $\mathcal{C}(t)$  is a *convex combination* of the control points

finally, we get the *convex hull property*,

- A Bézier curve lies in the convex hull of its control points

## 2.3 DeCasteljau Algorithm

In this section, we introduce the deCasteljau algorithm Algo. (2). The basic idea comes from the following remark; using the recurrence formulae Eq. (2.2), we have

$$\begin{aligned} \mathcal{C}(t) &= \sum_{k=0}^n \mathbf{P}_k B_k^n(t) = \sum_{k=0}^n \mathbf{P}_k ((1-t)B_k^{n-1}(t) + tB_{k-1}^{n-1}(t)) \\ &= \sum_{k=0}^{n-1} \mathbf{P}_k^1 B_k^{n-1}(t) && [\mathbf{P}_k^1 := (1-t)\mathbf{P}_k + t\mathbf{P}_{k+1}] \\ &= \sum_{k=0}^{n-2} \mathbf{P}_k^2 B_k^{n-2}(t) && [\mathbf{P}_k^2 := (1-t)\mathbf{P}_k^1 + t\mathbf{P}_{k+1}^1] \\ &= \dots \\ &= \sum_{k=0}^0 \mathbf{P}_k^n B_k^0(t) && [\mathbf{P}_k^n := (1-t)\mathbf{P}_k^{n-1} + t\mathbf{P}_{k+1}^{n-1}] \\ &= \mathbf{P}_0^n \end{aligned}$$

---

**Algorithm 2:** DECASTELJAU: Evaluation of a Bézier curve, defined by its control points  $\mathbf{P}$  at  $x$ .

---

**Input :**  $\mathbf{P}, x$   
**Output :** *value*

```

1  $n \leftarrow \text{len}(\mathbf{P}) - 1$ 
2  $\mathbf{Q} \leftarrow \mathbf{P}$ 
3 for  $j \leftarrow 0$  to  $n$  do
4   | for  $i \leftarrow 0$  to  $n - j$  do
5   |   |  $Q[i] \leftarrow (1 - x)Q[i] + xQ[i + 1]$ 
6   |   end
7 end
8 return  $Q[0]$ 

```

---

## 2.4 Conversion from/to monomial form

### Conversion from monomial form

Let us consider a monomial representation of a curve

$$\mathcal{C}(t) = \sum_{k=0}^n \mathbf{Q}_k t^k$$

Since

$$\begin{aligned} t^k &= \frac{1}{\binom{n}{k}} \binom{n}{k} t^k (1-t+t)^{n-k} = \frac{1}{\binom{n}{k}} \left( \sum_{i=0}^{n-k} \binom{n}{k} \binom{n-k}{n-k-i} t^{i+k} (1-t)^{n-k-i} \right) \\ &= \frac{1}{\binom{n}{k}} \left( \sum_{i=0}^{n-k} \binom{k+i}{k} B_{k+i}^n(t) \right) = \frac{1}{\binom{n}{k}} \left( \sum_{j=0}^n \binom{j}{k} B_j^n(t) \right) \end{aligned}$$

we get,

$$\mathcal{C}(t) = \sum_{k=0}^n \mathbf{Q}_k t^k = \sum_{k=0}^n \mathbf{Q}_k \frac{1}{\binom{n}{k}} \left( \sum_{j=0}^n \binom{j}{k} B_j^n(t) \right)$$

hence,

$$\mathcal{C}(t) = \sum_{j=0}^n \left( \sum_{k=0}^n \frac{\binom{j}{k}}{\binom{n}{k}} \mathbf{Q}_k \right) B_j^n(t)$$

which is a Bézier curve of the form

$$\mathcal{C}(t) = \sum_{k=0}^n \mathbf{P}_k B_k^n(t)$$

with

$$\mathbf{P}_k = \sum_{j=0}^n \frac{\binom{j}{k}}{\binom{n}{k}} \mathbf{Q}_k$$

**R** If  $\mathbf{Q}_k = 0$  for all  $k \geq 2$ , then the control points are given by  $\mathbf{P}_k = \mathbf{Q}_0 + k\mathbf{Q}_1$ .

**Proposition 2.4.1 — Linear Precision.** Conversely, if the  $n + 1$  control points  $\mathbf{P}$  lie equidistantly on a line, then  $\mathcal{C}$  is a linear polynomial, which can be written as  $\mathcal{C}(t) = (1 - t)\mathbf{P}_0 + t\mathbf{P}_n$ .

### Conversion to monomial form

Given a Bézier curve  $\mathcal{C}(t) = \sum_{k=0}^n \mathbf{P}_k B_k^n(t)$ , we can derive its monomial representation using the Taylor expansion,

$$\begin{aligned} \mathcal{C}(t) &= \sum_{k=0}^n \frac{1}{k!} \mathcal{C}^{(k)}(0) t^k \\ &= \sum_{k=0}^n \binom{n}{k} \Delta^k \mathbf{P}_0 t^k \end{aligned}$$

## 2.5 Rational Bézier curves

We first introduce the notion of Rational Bernstein polynomials defined as, and  $w_i > 0, \forall i \in [0, n]$

$$R_k^n(t) := \frac{w_k B_k^n(t)}{\sum_{i=0}^n w_i B_i^n(t)} \quad (2.10)$$

### Properties of Rational Bernstein polynomials

- $R_k^n(x) \geq 0$ , for all  $k \in [0, n]$  and  $x \in [0, 1]$  [positivity]
- $\sum_{k=0}^n R_k^n(x) = 1$ , for all  $x \in [0, 1]$  [partition of unity]
- $R_0^n(0) = R_n^n(1) = 1$
- $R_k^n$  has exactly one maximum on the interval  $[0, 1]$ ,
- Bernstein polynomials are Rational Bernstein polynomials when all weights are equal

#### Definition 2.5.1 — Rational Bézier curve.

$$\mathcal{C}(t) = \sum_{k=0}^n \mathbf{P}_k R_k^n(t) \quad (2.11)$$

where  $R_k^n(t) := \frac{w_k B_k^n(t)}{\sum_{i=0}^n w_i B_i^n(t)}$  and  $w_i > 0, \forall i \in [0, n]$

### Examples

**Example 1.** We consider the quadratic Rational Bézier curve,

where the weights are given by  $w_0 = 1$ ,  $w_1 = 1$  and  $w_2 = 2$  and the control points are

$$\mathbf{P}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{P}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \mathbf{P}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This leads to a representation of the circular arc, using the parametric form  $\mathcal{C}(t) = \begin{pmatrix} \frac{1-t^2}{1+t^2} \\ \frac{2t}{1+t^2} \end{pmatrix}$ .

## 2.6 Composite Bézier curves

A composite Bézier curve is a piecewise Bézier curve that is at least continuous at the interpolant control points. An example is given in Fig. 2.9. The global curve inherits all the nice properties of Bézier curves however, it has a major drawback: when moving an interpolant control point, we must ensure that the associated regularity is not broken, unless it is what we want. This means that a good representation of a curve needs to have the locality control through control points, but also the control points must be associated to some given regularity. Since Bernstein polynomials are defined locally on the interval  $[0, 1]$ , and they do not encode any global regularity of the curve, we need another representation, in terms of other functions, while keeping in mind that these functions must be polynomials on every sub-interval. Such representation is provided by the Schoenberg spaces, for which B-Splines form a basis. In the next chapter, we will first introduce the uniform Splines, also known as Cardinal Splines, show their advantages and limitations, then we will introduce the general definition of B-Splines.

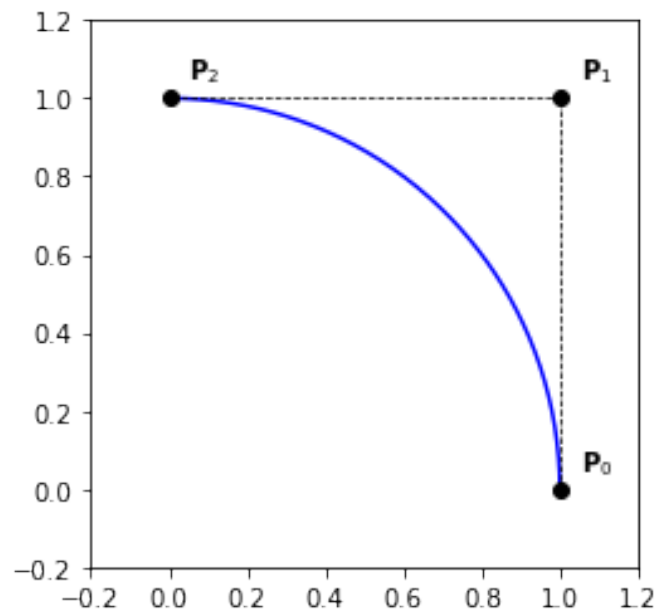


Figure 2.8: Circular arc using quadratic Rational Bézier curve.

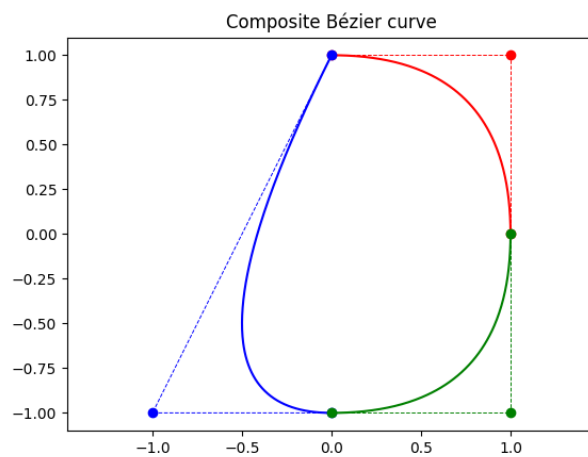


Figure 2.9: A composite Bézier curve using 3 quadratic Bézier curves.



## 2.7 Problems

**Exercise 2.1** Show all Bernstein properties. ■

**Exercise 2.2** Consider a cubic Bézier curve  $\mathcal{C}$  in  $\mathbb{R}^2$  with the following control points:

$$\mathbf{P}_0 = (0, 6), \quad \mathbf{P}_1 = (3, 6), \quad \mathbf{P}_2 = (6, 3), \quad \mathbf{P}_3 = (6, 0)$$

Compute the point  $\mathcal{C}(\frac{1}{3})$  using the deCasteljau algorithm. Compute the same point by Eqs. (??) and (??), meaning, evaluate the basis functions at  $u = \frac{1}{3}$  then multiply by the associated control points. ■

**Exercise 2.3** The Bernstein operator  $\mathcal{B}$  assigns to a function  $f$  on  $[0, 1]$  the polynomial

$$\mathcal{B}[f](t) := \sum_{i=0}^n f\left(\frac{i}{n}\right) B_i^n(t)$$

Show that if  $f$  is a polynomial of degree  $m \leq n$ , then  $\mathcal{B}[f]$  is also a polynomial of degree  $m$ . ■

**Exercise 2.4** Show that a planar cubic Bézier curve has a cusp if  $\mathbf{P}_3$  lies on the parabola

$$t \rightarrow (\mathbf{P}_0 + \mathbf{P}_1 - \mathbf{P}_2) B_0^2(t) + \mathbf{P}_1 B_1^2(t) + \mathbf{P}_2 B_2^2(t)$$

**Exercise 2.5** For which choices of  $\mathbf{P}_3$  does a planar cubic Bézier curve have a loop? ■

**Exercise 2.6** We consider a Bézier curve  $\mathcal{C}(t) = \sum_{k=0}^n \mathbf{P}_k B_k^n(t)$ .

1. Show that using the definition of Bernstein polynomials, we can write  $\mathcal{C}$  as

$$\mathcal{C}(t) = (1-t)^n \left( \sum_{i=0}^n \mathbf{P}_k \binom{n}{k} \left( \frac{t}{1-t} \right)^k \right)$$

2. Write a modified version of Horner algorithm based on the previous formula and compute its arithmetic complexity.
3. What happens when  $t$  is close to 1? How to adapt the previous algorithm in this case? ■

**Exercise 2.7** . ■

**Exercise 2.8** . ■

**Exercise 2.9** . ■

**Exercise 2.10** . ■

**Exercise 2.11** .



**Exercise 2.12** .



## 3. B-Splines

In general, one needs to design curves with given regularities at some specific locations. For this purpose, rather than considering piecewise Bézier curves with additional constraints on the control points, we can enforce these constraints on a given space of functions. The Schoenberg spaces [1946] are perfect for this purpose. Before giving more details about Schoenberg spaces, which will be covered in the Approximation Theory part, let us for the moment, use a formal definition, as the following: given a subdivision  $\{x_0 < x_1 < \dots < x_r\}$  of the interval  $I = [x_0, x_r]$ , the Schoenberg space is the space of piecewise polynomials of degree  $p$ , on the interval  $I$  and given regularities  $\{k_1, k_2, \dots, k_{r-1}\}$  at the internal points  $\{x_1, x_2, \dots, x_{r-1}\}$ .

Given  $m$  and  $p$  as natural numbers, let us consider a sequence of non decreasing real numbers  $T = \{t_i\}_{0 \leq i \leq m}$ .  $T$  is called **knots sequence**. From a knots sequence, we can generate a B-splines family using the recurrence formula 3.1.

**Definition 3.0.1 — B-Splines using Cox-DeBoor Formula.** The  $j$ -th B-spline of degree  $p$  is defined by the recurrence relation:

$$N_j^p = \frac{t - t_j}{t_{j+p} - t_j} N_j^{p-1} + \frac{t_{j+p+1} - t}{t_{j+p+1} - t_{j+1}} N_{j+1}^{p-1}, \quad (3.1)$$

where

$$N_j^0(t) = \chi_{[t_j, t_{j+1}]}(t)$$

for  $0 \leq j \leq m - p - 1$ .

**R** When working with Bernstein polynomials, we introduced the convention  $B_i^n = 0$  for all  $i < 0$  or  $i > n$ . For B-Splines, we have a similar convention  $N_j^p = 0$  if  $j < 0$  or  $j > n$ . In addition, we also assume  $\frac{0}{0} = 0$ , when using the formula 3.1 and  $N_j^0 = 0$  if  $t_j = t_{j+1}$ .

**R**  $k = p + 1$  is called the order of the B-Splines.

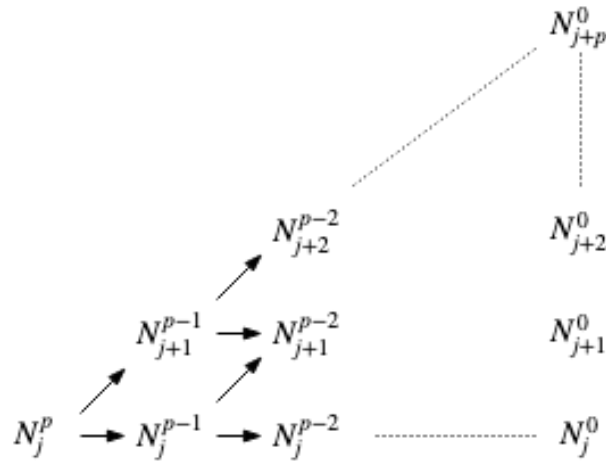


Figure 3.1: General evaluation triangular diagram for a B-Spline.

In figure Fig. 3.2, we plot the quadratic B-Splines functions associated to the knots sequence  $\{0, 0, 0, 1, 2, 2, 3, 4, 5, 5, 5\}$ . As we can see, there are 8 B-Splines functions and we recover the interpolation property at the first and last knot. In addition, while all the B-Splines are *smooth* at the internal knots,  $N_3^2$  is only  $\mathcal{C}^0$  at the grid point 2.

Now if we look for the cubic B-Splines that are associated to the same knots sequence  $\{0, 0, 0, 1, 2, 2, 3, 4, 5, 5, 5\}$ , we see that the interpolation property is no longer true 3.3.

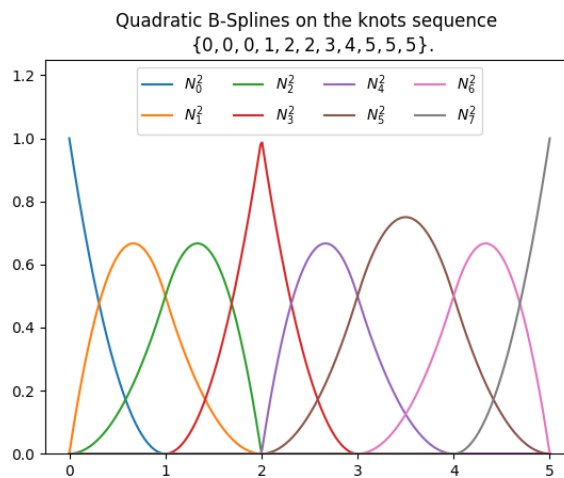


Figure 3.2: Quadratic B-Splines generated using the knots sequence  $\{0, 0, 0, 1, 2, 2, 3, 4, 5, 5, 5\}$ .

### 3.1 Knot vector families

There are two kind of **knot vectors**, called **clamped** and **unclamped**. Both families contains uniform and non-uniform sequences. In figure Fig. ??, we show examples of clamped uniform knots. Fig. ??, we show examples of clamped non-uniform knots, while figures Fig. ?? and Fig. ?? are examples of unclamped knots sequences for uniform and non-uniform cases, respectively.

The special case of clamped sequences, where the first and the last knots are repeated  $p + 1$  times exactly is known as **open knot vector**.

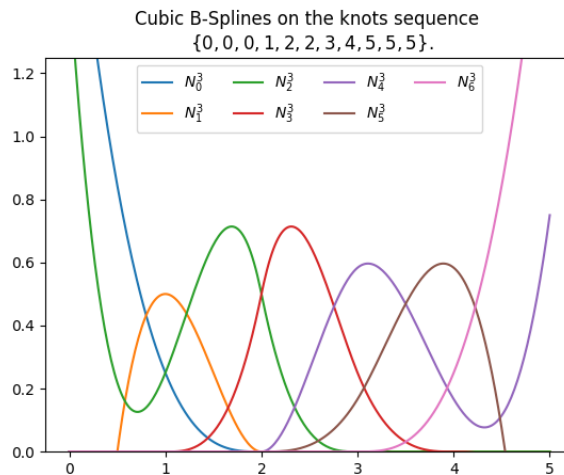


Figure 3.3: Cubic B-Splines generated using the knots sequence  $\{0, 0, 0, 1, 2, 2, 3, 4, 5, 5, 5\}$ .

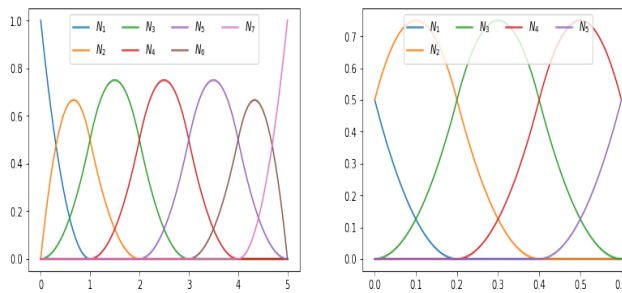


Figure 3.4: **Clamped uniform knots** (left) quadratic B-Splines using  $T = \{0, 0, 0, 1, 2, 3, 4, 5, 5, 5\}$ , (right) quadratic B-Splines using  $T = \{-0.2, -0.2, 0.0, 0.2, 0.4, 0.6, 0.8, 0.8\}$ .

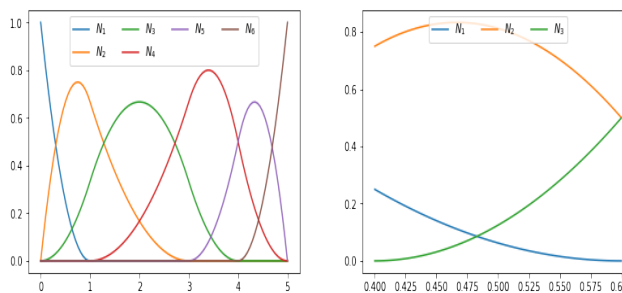


Figure 3.5: **Clamped non-uniform knots** (left) quadratic B-Splines using  $T = \{0, 0, 0, 1, 3, 4, 5, 5, 5\}$ , (right) quadratic B-Splines using  $T = \{-0.2, -0.2, 0.4, 0.6, 0.8, 0.8\}$ .

### 3.2 Examples

Before stating some B-Splines properties, let us take a look to some examples.

**Example 1.** Let  $T = \{0, 1, 2\}$  and  $p = 1$ .

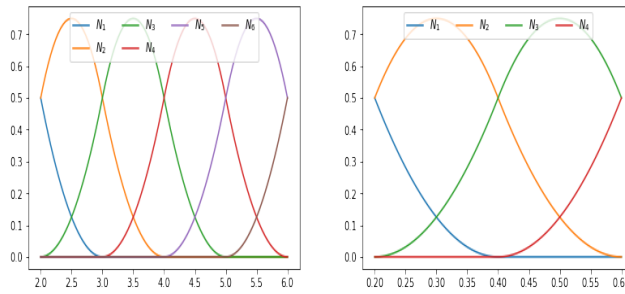


Figure 3.6: **Unclamped uniform knots** (left) quadratic B-Splines using  $T = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$ , (right) quadratic B-Splines using  $T = \{-0.2, 0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ .

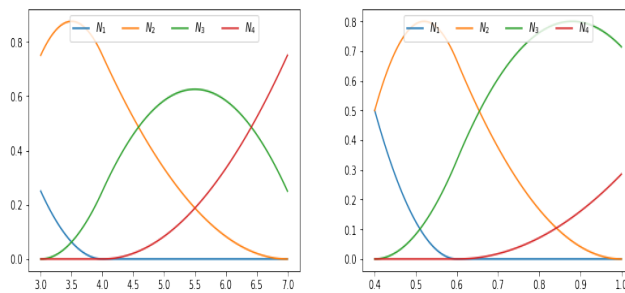


Figure 3.7: **Unclamped non-uniform knots** (left) quadratic B-Splines using  $T = \{0, 0, 3, 4, 7, 8, 9\}$ , (right) quadratic B-Splines using  $T = \{-0.2, 0.2, 0.4, 0.6, 1.0, 2.0, 2.5\}$ .

We first compute the B-Splines of degrees 0

$$N_0^0 = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$N_1^0 = \begin{cases} 1, & 1 \leq t < 2 \\ 0, & \text{otherwise} \end{cases}$$

The B-Splines of degree 1 is

$$N_0^1 = \frac{t-0}{1-0}N_0^0 + \frac{2-t}{2-1}N_1^0$$

which yields

$$N_0^1 = \begin{cases} t, & 0 \leq t < 1 \\ 2-t, & 1 \leq t < 2 \\ 0, & \text{otherwise} \end{cases}$$

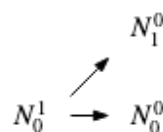


Figure 3.8: Evaluation diagram for examples 1, 2 and 3.

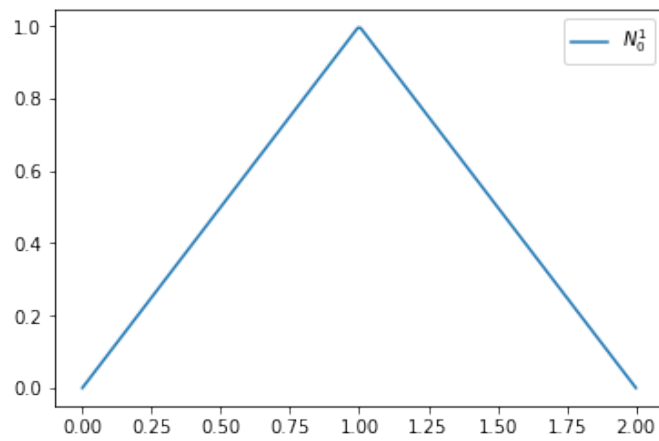


Figure 3.9: Linear B-Spline associated to the knot sequence  $T = \{0, 1, 2\}$ .

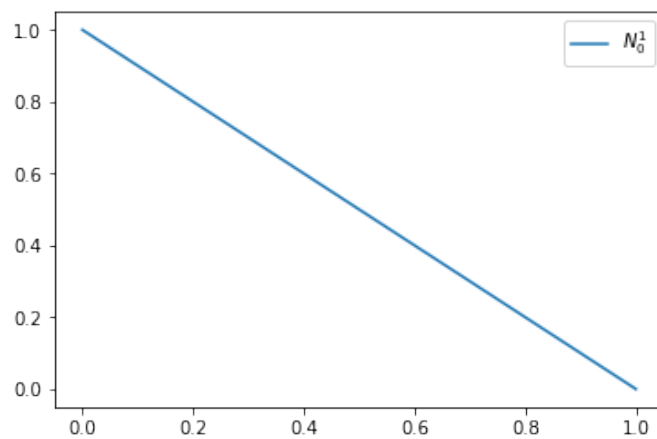


Figure 3.10: Linear B-Spline associated to the knot sequence  $T = \{0, 0, 1\}$ .

**Example 2.** Let  $T = \{0, 0, 1\}$  and  $p = 1$ .

We first compute the B-Splines of degrees 0

$$N_0^0 = 0, \quad \forall t \in \mathbb{R}$$

$$N_1^0 = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

The B-Splines of degree 1 is

$$N_0^1 = \frac{t-0}{0-0}N_0^0 + \frac{1-t}{1-0}N_1^0$$

which yields

$$N_0^1 = \begin{cases} 1-t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

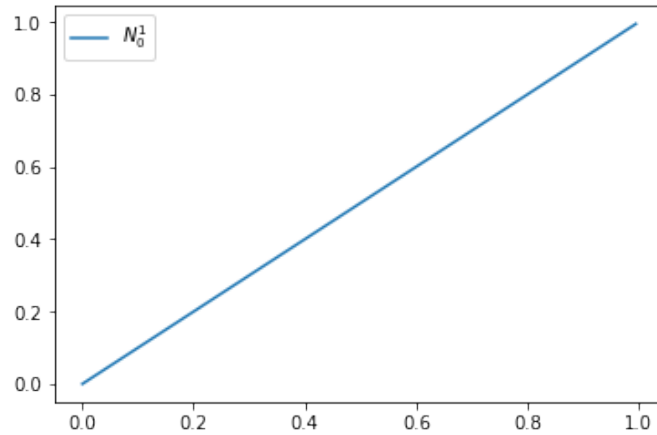


Figure 3.11: Linear B-Spline associated to the knot sequence  $T = \{0, 1, 1\}$ .

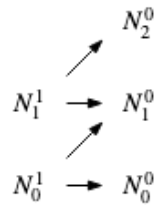


Figure 3.12: Evaluation diagram for examples 4 and 5.

**Example 3.** Let  $T = \{0, 1, 1\}$  and  $p = 1$ .

We first compute the B-Splines of degrees 0

$$N_0^0 = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$N_1^0 = 0, \quad \forall t \in \mathbb{R}$$

The B-Splines of degree 1 is

$$N_0^1 = \frac{t-0}{1-0}N_1^0 + \frac{1-t}{1-1}N_2^0$$

which yields

$$N_0^1 = \begin{cases} t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

**Example 4.** Let  $T = \{0, 0, 1, 1\}$  and  $p = 1$ .

In the examples 2 and 3 we computed the linear B-Splines associated to the knots sequences  $\{0, 0, 1\}$  and  $\{0, 1, 1\}$ . Then we get immediately that the knots sequences  $\{0, 0, 1, 1\}$  will exactly generate the same B-Splines, *i.e.*

$$N_0^1 = \begin{cases} 1-t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$N_1^1 = \begin{cases} t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$



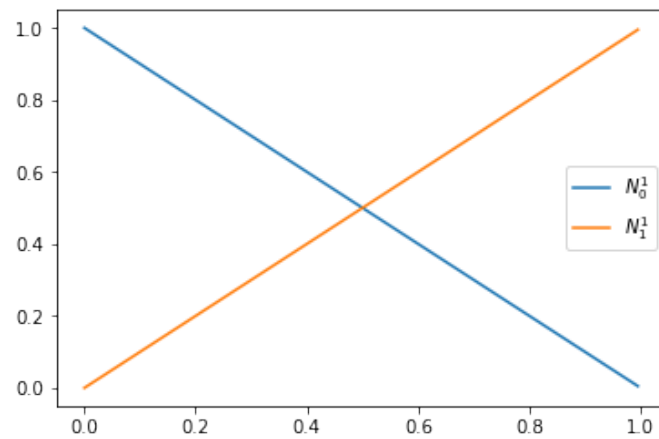


Figure 3.13: Linear B-Splines associated to the knots sequence  $T = \{0, 0, 1, 1\}$ .

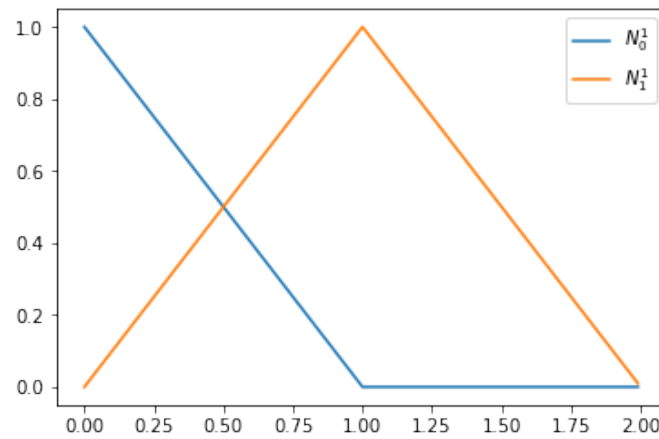


Figure 3.14: Linear B-Splines associated to the knot sequence  $T = \{0, 0, 1, 2\}$ .

**Example 5.** Let  $T = \{0, 0, 1, 2\}$  and  $p = 1$ .

In the examples 2 and 1 we computed the linear B-Splines associated to the knots sequences  $\{0, 0, 1\}$  and  $\{0, 1, 2\}$ . Then we get immediately that the knots sequences  $\{0, 0, 1, 2\}$  will exactly generate the same B-Splines, *i.e.*

$$N_0^1 = \begin{cases} 1-t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$N_1^1 = \begin{cases} t, & 0 \leq t < 1 \\ 2-t, & 1 \leq t < 2 \\ 0, & \text{otherwise} \end{cases}$$

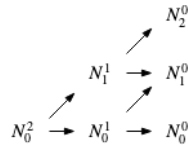
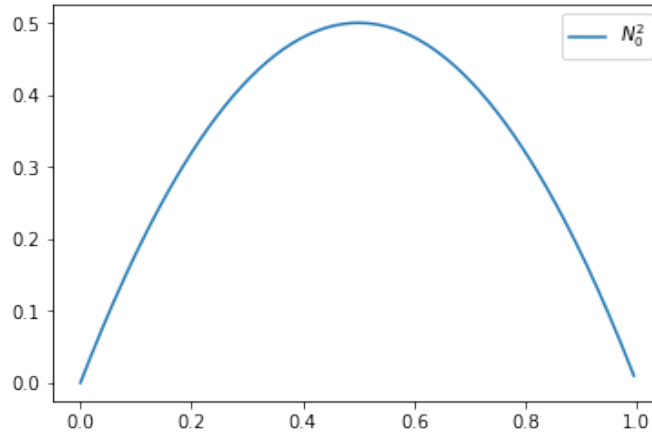


Figure 3.15: Evaluation diagram for examples 6 and 7.

Figure 3.16: Quadratic B-Splines associated to the knots sequence  $T = \{0, 0, 1, 1\}$ .

**Example 6.** Let  $T = \{0, 0, 1, 1\}$  and  $p = 2$ .

The linear B-Splines were computed in the previous example. Then for degree 2, we get

$$N_0^2 = \frac{t-0}{1-0}N_0^1 + \frac{1-t}{1-0}N_1^1$$

which leads to

$$N_0^2 = \begin{cases} 2(1-t)t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

**Example 7.** Let  $T = \{0, 0, 1, 2\}$  and  $p = 2$ .

The linear B-Splines were computed in the examples 2 and 1. Then for degree 2, we get

$$N_0^2 = \frac{t-0}{1-0}N_0^1 + \frac{2-t}{2-0}N_1^1$$

which yields

$$N_0^2 = \begin{cases} 2t - \frac{3}{2}t^2, & 0 \leq t < 1 \\ \frac{1}{2}(2-t)^2, & 1 \leq t < 2 \\ 0, & \text{otherwise} \end{cases}$$

**Example 8.** Let  $T = \{0, 0, 1, 2, 3, 3\}$  and  $p = 1$ .

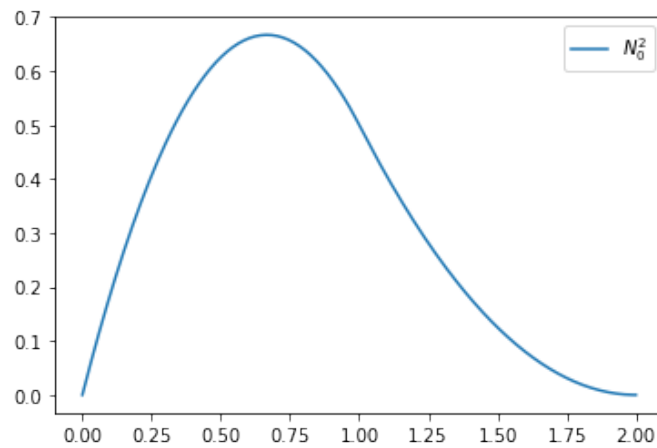


Figure 3.17: Quadratic B-Splines associated to the knots sequence  $T = \{0, 0, 1, 2\}$ .

We first compute the B-Splines of degrees 0

$$N_0^0 = 0, \quad \forall t \in \mathbb{R}$$

$$N_1^0 = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$N_2^0 = \begin{cases} 1, & 1 \leq t < 2 \\ 0, & \text{otherwise} \end{cases}$$

$$N_3^0 = \begin{cases} 1, & 2 \leq t < 3 \\ 0, & \text{otherwise} \end{cases}$$

$$N_4^0 = 0, \quad \forall t \in \mathbb{R}$$

The B-Splines of degree 1 are

$$N_0^1 = \frac{t-0}{0-0}N_0^0 + \frac{1-t}{1-0}N_1^0$$

$$N_1^1 = \frac{t-0}{1-0}N_1^0 + \frac{2-t}{2-1}N_2^0$$

$$N_2^1 = \frac{t-1}{2-1}N_2^0 + \frac{3-t}{3-2}N_3^0$$

$$N_3^1 = \frac{t-2}{3-2}N_3^0 + \frac{3-t}{3-3}N_4^0$$

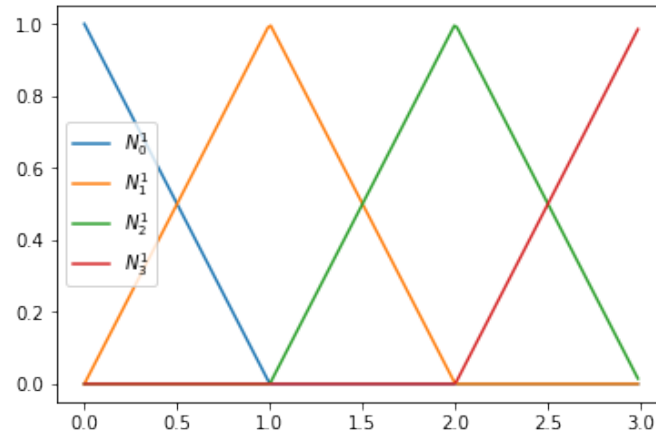


Figure 3.18: Linear B-Splines associated to the knots sequence  $T = \{0, 0, 1, 2, 3, 3\}$ .

which gives

$$N_0^1 = \begin{cases} 1-t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$N_1^1 = \begin{cases} t, & 0 \leq t < 1 \\ 2-t, & 1 \leq t < 2 \\ 0, & \text{otherwise} \end{cases}$$

$$N_2^1 = \begin{cases} t-1, & 1 \leq t < 2 \\ 3-t, & 2 \leq t < 3 \\ 0, & \text{otherwise} \end{cases}$$

$$N_3^1 = \begin{cases} t-2, & 2 \leq t < 3 \\ 0, & \text{otherwise} \end{cases}$$

**Example 9.** Let  $T = \{0, 0, 0, 1, 1, 1\}$  and  $p = 2$ .

We first compute the B-Splines of degrees 0

$$N_0^0 = N_1^0 = 0, \quad \forall t \in \mathbb{R}$$

$$N_2^0 = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$N_3^0 = N_4^0 = 0, \quad \forall t \in \mathbb{R}$$

The B-Splines of degree 1 are

$$N_0^1 = \frac{t-0}{0-0}N_0^0 + \frac{0-t}{0-0}N_1^0$$

$$N_1^1 = \frac{t-0}{0-0}N_1^0 + \frac{1-t}{1-0}N_2^0$$

$$N_2^1 = \frac{t-0}{1-0}N_2^0 + \frac{1-t}{1-1}N_3^0$$

$$N_3^1 = \frac{t-1}{1-1}N_3^0 + \frac{1-t}{1-1}N_4^0$$

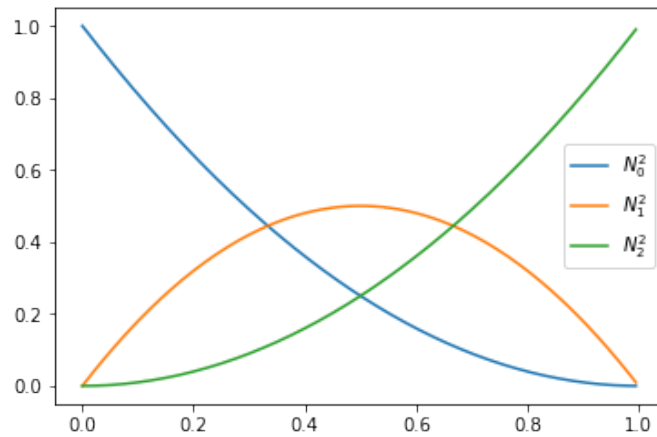


Figure 3.19: Linear B-Splines associated to the knots sequence  $T = \{0, 0, 0, 1, 1, 1\}$ .

which leads to

$$\begin{aligned}
 N_0^1 &= 0, \quad \forall t \in \mathbb{R} \\
 N_1^1 &= \begin{cases} 1-t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \\
 N_2^1 &= \begin{cases} t, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \\
 N_3^1 &= 0, \quad \forall t \in \mathbb{R}
 \end{aligned}$$

The B-Splines of degree 2 are

$$\begin{aligned}
 N_0^2 &= \frac{t-0}{0-0}N_0^1 + \frac{1-t}{1-0}N_1^1 \\
 N_1^2 &= \frac{t-0}{1-0}N_1^1 + \frac{1-t}{1-0}N_2^1 \\
 N_2^2 &= \frac{t-0}{1-0}N_2^1 + \frac{1-t}{1-1}N_3^1
 \end{aligned}$$

which leads to

$$\begin{aligned}
 N_0^2 &= \begin{cases} (1-t)^2, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \\
 N_1^2 &= \begin{cases} 2t(1-t), & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \\
 N_2^2 &= \begin{cases} t^2, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}
 \end{aligned}$$

In this case, we recover the Bernstein polynomials of degree 2.

**Exercise 3.1** Compute the B-Splines generated by the knots sequences:

- $T = \{0, 0, 0, 1, 2, 2, 3, 4, 5, 5, 5\}$  with  $p = 2$
- $T = \{0, 1, 2, 2, 3, 4, 5, 5, 6\}$  with  $p = 2$

### 3.3 B-Splines properties

B-Splines have many interesting properties that are listed below. In general, most of the proofs are done by induction on the B-Spline degree.

**Proposition 3.3.1** B-splines are piecewise polynomial of degree  $p$

*Proof.* We proceed by induction. When  $p = 0$ , B-Splines are either the 0 or the characteristic functions, which are constant polynomials. For  $p = 1$ , the Cox-DeBoor formula shows that the generated B-Splines are piecewise linear functions, because of the weights that multiply the constant functions (B-Splines of degree  $p = 0$ ).

Now assume that B-Splines of degree  $p - 1$  are piecewise polynomials of degree  $p - 1$ . Again using the Cox-DeBoor formula, we see that the degree will be increased by 1 because of the linear weights, on each sub-interval. This completes the proof. ■

**R** The B-splines functions associated to open knots sequences without internal knots, *i.e.* the length of the knots sequence is exactly  $2p + 2$ , are the Bernstein polynomials of degree  $p$ .

**Proposition 3.3.2 — compact support.**  $N_j^p(t) = 0$  for all  $t \notin [t_j, t_{j+p+1})$ .

*Proof.* By taking a look at the triangular diagram 3.1 for the evaluation of the  $j$ th B-Spline of degree  $p$ . We see that  $N_j^p$  depends only on the B-Splines  $\{N_j^0, N_{j+1}^0, \dots, N_{j+p}^0\}$ . The latter are exactly the characteristic functions on the intervals  $[t_i, t_{i+1}), \forall i \in \{j, j+1, \dots, j+p\}$ . Therefore,  $N_j^p(t) = 0$ , for all  $t \notin [t_j, t_{j+p+1})$ . ■

**Proposition 3.3.3** If  $t \in [t_j, t_{j+1})$ , then only the B-splines  $\{N_{j-p}^p, \dots, N_j^p\}$  are non vanishing at  $t$ .

*Proof.* Using the last proposition, the support of  $N_i^p$  is defined as  $\text{supp}(N_i^p) = [t_i, t_{i+p+1})$  for every  $i \in \{j-p, \dots, j\}$ . Therefore, on the interval  $[t_j, t_{j+1}) = \bigcap_{j-p \leq i \leq j} \text{supp}(N_i^p)$ , only the B-Splines  $\{N_{j-p}^p, \dots, N_j^p\}$  are non-zeros. ■

**Proposition 3.3.4 — non-negativity.**  $N_j^p(t) \geq 0, \quad \forall t \in [t_j, t_{j+p+1})$

*Proof.* This can be proved by induction. It is true for  $p = 0$ , since all B-Splines are either 0 or the characteristic function. Assume the property holds for  $p - 1$ , using the Cox-DeBoor formula we have:

$$N_j^p = \frac{x - t_j}{t_{j+p} - t_j} N_j^{p-1} + \frac{t_{j+p+1} - x}{t_{j+p+1} - t_{j+1}} N_{j+1}^{p-1}$$

since,  $t \in [t_j, t_{j+p+1})$ , we have  $\frac{x - t_j}{t_{j+p} - t_j} \geq 0$  and  $\frac{t_{j+p+1} - x}{t_{j+p+1} - t_{j+1}} \geq 0$  whenever  $t_{j+p} > t_j$  and  $t_{j+p+1} > t_{j+1}$ , otherwise these terms are identically equal to 0, by convention.

On the other hand,  $N_j^{p-1} \geq 0$  and  $N_{j+1}^{p-1} \geq 0$ . Therefore there linear combination is also positive. ■

**Proposition 3.3.5 — Partition of unity.**  $\sum N_i^p(t) = 1, \forall t \in \mathbb{R}$

**R** The previous sum  $\sum N_i^p(t)$  has a meaning since for every  $t \in \mathbb{R}$ , only  $p + 1$  B-Splines are non-vanishing.

*Proof.* We prove this result by induction again. For  $p = 0$ , the property is true. Now let us assume it is true for  $p - 1$ .

Let  $j$  denotes the span of  $t$ . We know that  $\sum N_i^p(t) = \sum_{i=j-p}^j N_i^p(t)$ . Now, using the Cox-DeBoor formula, we have

$$\sum_{i=j-p}^j N_i^p(t) = \sum_{i=j-p}^j \left( \frac{t-t_i}{t_{i+p}-t_i} N_i^{p-1}(t) + \frac{t_{i+p+1}-t}{t_{i+p+1}-t_{i+1}} N_{i+1}^{p-1}(t) \right)$$

by changing the summation index in the second part, we get

$$\sum_{i=j-p}^j N_i^p(t) = \sum_{i=j-p}^j \frac{t-t_i}{t_{i+p}-t_i} N_i^{p-1}(t) + \sum_{i=j-p+1}^{j+1} \frac{t_{i+p}-t}{t_{i+p}-t_i} N_i^{p-1}(t)$$

On the other hand, we know that the only B-Splines of degree  $p - 1$  that are non-vanishing on the interval  $[t_j, t_{j+1})$  are  $\{N_{j-p+1}^{p-1}, \dots, N_j^{p-1}\}$ , meaning that  $N_{j-p}^{p-1} = N_{j+1}^{p-1} = 0$ . Therefore,

$$\sum_{i=j-p}^j N_i^p(t) = \sum_{i=j-p+1}^j \left( \frac{t-t_i}{t_{i+p}-t_i} + \frac{t_{i+p}-t}{t_{i+p}-t_i} \right) N_i^{p-1}(t)$$

hence,

$$\sum_{i=j-p}^j N_i^p(t) = \sum_{i=j-p+1}^j N_i^{p-1}(t)$$

Finally, we use the induction assumption to complete the proof. ■

**R** For the sake of simplicity, we shall avoid using summation indices on linear expansion of B-Splines.

#### Lemma 1

$$\sum a_j N_j^p = \sum \left( \frac{t_{j+p}-t}{t_{j+p}-t_j} a_{j-1} + \frac{t-t_j}{t_{j+p}-t_j} a_j \right) N_j^{p-1} \quad (3.2)$$

*Proof.* Using the Cox-DeBoor formula, we have

$$\sum a_j N_j^p = \sum \left( a_j \frac{t-t_j}{t_{j+p}-t_j} N_j^{p-1} + a_j \frac{t_{j+p+1}-t}{t_{j+p+1}-t_{j+1}} N_{j+1}^{p-1} \right)$$

By changing the summation index in the second part of the right hand side, we get

$$\sum a_j N_j^p = \sum a_j \frac{t-t_j}{t_{j+p}-t_j} N_j^{p-1} + \sum a_{j-1} \frac{t_{j+p}-t}{t_{j+p}-t_j} N_j^{p-1}$$

which yields

$$\sum a_j N_j^p = \sum \left( a_j \frac{t-t_j}{t_{j+p}-t_j} + a_{j-1} \frac{t_{j+p}-t}{t_{j+p}-t_j} \right) N_j^{p-1}$$

■

**Proposition 3.3.6 — Marsden's identity.**

$$(x - y)^p = \sum_j \rho_j^p(y) N_j^p(x) \quad (3.3)$$

where

$$\rho_j^p(y) = \prod_{i=j+1}^{j+p} (t_i - y) \quad (3.4)$$

*Proof.* Setting  $a_j := \rho_j^p(y)$  in the lemma 1, we get

$$\sum \rho_j^p(y) N_j^p(x) = \sum \left( \frac{t_{j+p} - x}{t_{j+p} - t_j} \rho_{j-1}^p(y) + \frac{x - t_j}{t_{j+p} - t_j} \rho_j^p(y) \right) N_j^{p-1}(x)$$

on the other hand, by introducing  $\alpha_j^{p-1} = \prod_{i=j+1}^{j+p-1} (t_i - y)$ , we have

$$\begin{aligned} \frac{t_{j+p} - x}{t_{j+p} - t_j} \rho_{j-1}^p(y) + \frac{x - t_j}{t_{j+p} - t_j} \rho_j^p(y) &= \alpha_j^{p-1} \left( \frac{t_{j+p} - x}{t_{j+p} - t_j} (t_j - y) + \frac{x - t_j}{t_{j+p} - t_j} (t_{j+p} - y) \right) \\ &= (x - y) \rho_j^{p-1}(y) \end{aligned}$$

Therefore,

$$\sum \rho_j^p(y) N_j^p(x) = (x - y) \sum \rho_j^{p-1}(y) N_j^{p-1}(x)$$

By repeating the same computation  $p - 1$  times (one can also use an induction for this), we get the desired result. ■

Thanks to Marsden's identity and the proposition 3.3.3, we now are able to prove the local linear independence of the B-Splines basis functions.

**Proposition 3.3.7 — Local linear independence.** On each interval  $[t_j, t_{j+1})$ , the B-Splines are linearly independent.

*Proof.* On the interval  $[t_j, t_{j+1})$ , there are only  $p + 1$  non-vanishing B-Splines. On the other hand, these B-Splines reproduce polynomials, thanks to the Marsden's identity. Hence,  $\{N_{j-p}^p, \dots, N_j^p\}$  are a basis for polynomials of degree at most  $p$ , and are therefore, linearly independent. ■

**R** We will see in the Approximation theory part, that the B-Splines family has also a global linear independence property.

### 3.4 Derivatives of B-Splines

The derivative of B-Splines can be computed recursively by deriving the formula 3.1, which gives

**Proposition 3.4.1**

$$N_j^{p'} = \frac{P}{t_{j+p} - t_j} N_j^{p-1} - \frac{P}{t_{j+p+1} - t_{j+1}} N_{j+1}^{p-1} \quad (3.5)$$



*Proof.* We prove this by induction on  $p$ .

For  $p = 1$ ,  $N_j^{p-1}$  and  $N_{j+1}^{p-1}$  are either 0 or 1, then  $N_j^{p'}$  is either  $\frac{1}{t_{j+1}-t_j}$  or  $-\frac{1}{t_{j+2}-t_{j+1}}$ .

Now assume that 3.5, is true for  $p - 1$ , where  $p > 1$ . Deriving the Cox-DeBoor formula leads to

$$\begin{aligned} N_j^{p'} &= \frac{1}{t_{j+p}-t_j} N_j^{p-1} + \frac{t-t_j}{t_{j+p}-t_j} N_j^{p-1'} \\ &\quad - \frac{1}{t_{j+p+1}-t_{j+1}} N_{j+1}^{p-1} + \frac{t_{j+p+1}-t}{t_{j+p+1}-t_{j+1}} N_{j+1}^{p-1'} \end{aligned}$$

Using the induction assumption on  $N_j^{p-1'}$  and  $N_{j+1}^{p-1'}$  and substituting in the last equation yields

$$\begin{aligned} N_j^{p'} &= \frac{1}{t_{j+p}-t_j} N_j^{p-1} - \frac{1}{t_{j+p+1}-t_{j+1}} N_{j+1}^{p-1} \\ &\quad + \frac{t-t_j}{t_{j+p}-t_j} \left( \frac{p-1}{t_{j+p-1}-t_j} N_j^{p-2} - \frac{p-1}{t_{j+p}-t_{j+1}} N_{j+1}^{p-2} \right) \\ &\quad + \frac{t_{j+p+1}-t}{t_{j+p+1}-t_{j+1}} \left( \frac{p-1}{t_{j+p}-t_{j+1}} N_{j+1}^{p-2} - \frac{p-1}{t_{j+p+1}-t_{j+2}} N_{j+2}^{p-2} \right) \\ &= \frac{1}{t_{j+p}-t_j} N_j^{p-1} - \frac{1}{t_{j+p+1}-t_{j+1}} N_{j+1}^{p-1} \\ &\quad + \frac{p-1}{t_{j+p}-t_j} \frac{t-t_j}{t_{j+p-1}-t_j} N_j^{p-2} \\ &\quad + \frac{p-1}{t_{j+p}-t_{j+1}} \left( \frac{t_{j+p+1}-t}{t_{j+p+1}-t_{j+1}} - \frac{t-t_j}{t_{j+p}-t_j} \right) N_{j+1}^{p-2} \\ &\quad - \frac{p-1}{t_{j+p+1}-t_{j+1}} \frac{t_{j+p+1}-t}{t_{j+p+1}-t_{j+2}} N_{j+2}^{p-2} \end{aligned}$$

Since,

$$\frac{t_{j+p+1}-t}{t_{j+p+1}-t_{j+1}} - \frac{t-t_j}{t_{j+p}-t_j} = \frac{t_{j+p}-t}{t_{j+p}-t_j} - \frac{t-t_{j+1}}{t_{j+p+1}-t_{j+1}}$$

we get

$$\begin{aligned} N_j^{p'} &= \frac{1}{t_{j+p}-t_j} N_j^{p-1} - \frac{1}{t_{j+p+1}-t_{j+1}} N_{j+1}^{p-1} \\ &\quad + \frac{p-1}{t_{j+p}-t_j} \left( \frac{t-t_j}{t_{j+p-1}-t_j} N_j^{p-2} + \frac{t_{j+p}-t}{t_{j+p}-t_j} N_{j+1}^{p-2} \right) \\ &\quad - \frac{p-1}{t_{j+p+1}-t_{j+1}} \left( \frac{t-t_{j+1}}{t_{j+p}-t_{j+1}} N_{j+1}^{p-2} + \frac{t_{j+p+1}-t}{t_{j+p+1}-t_{j+2}} N_{j+2}^{p-2} \right) \end{aligned}$$

By applying the Cox-DeBoor formula for  $N_j^{p-1}$  and  $N_{j+1}^{p-1}$  we get the desired result. ■

**Exercise 3.2** Show that the  $r$ -th derivative of  $N_j^p$  is given by

$$N_j^{p(r)} = \frac{p}{t_{j+p}-t_j} N_j^{p-1(r-1)} - \frac{p}{t_{j+p+1}-t_{j+1}} N_{j+1}^{p-1(r-1)}$$

### 3.5 Problems

TODO

## 4. Cardinal B-Splines

### 4.1 Cardinal B-Splines

Cardinal B-Splines play an important role in the approximation theory (multi-resolution approximation, ...). In the sequel, we shall give a definition of the Cardinal B-Spline using the convolution operator. Then, we will present some of the most important properties, at least needed when using uniform B-Splines in a Finite Elements method.

**Definition 4.1.1** A cardinal B-spline of zero degree, denoted by  $\phi_0$ , is the characteristic function over the interval  $[0, 1)$ , i.e.,

$$\phi_0(t) := \begin{cases} 1, & t \in [0, 1) \\ 0, & \text{otherwise} \end{cases} \quad (4.1)$$

A cardinal B-Spline of degree  $p$ ,  $p \in \mathbb{N}$ , denoted by  $\phi_p$ , is defined by convolution as

$$\phi_p(t) = (\phi_{p-1} * \phi_0)(t) = \int_{\mathbb{R}} \phi_{p-1}(t-s)\phi_0(s) ds \quad (4.2)$$

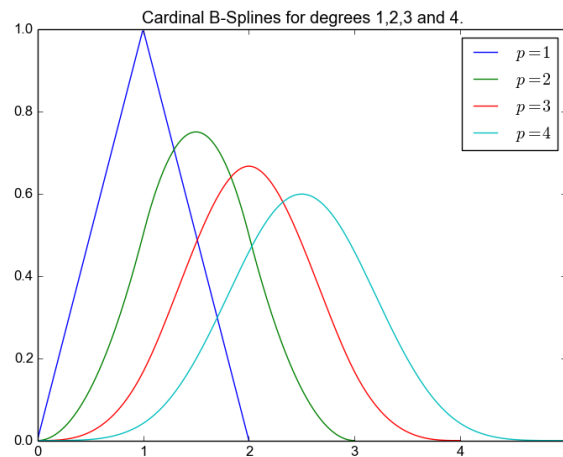


Figure 4.1: Cardinal B-Splines of degrees 1, 2, 3 and 4

**Example 1:**

When  $p = 1$ , it is easy to show that

$$\phi_1(t) := \begin{cases} t, & t \in [0, 1) \\ 2 - t, & t \in [1, 2) \\ 0, & \text{otherwise} \end{cases} \quad (4.3)$$

When  $p = 2$ , it is easy to show that

$$\phi_2(t) := \begin{cases} \frac{1}{2}t^2, & t \in [0, 1) \\ \frac{1}{2} + (t-1) - (t-1)^2, & t \in [1, 2) \\ \frac{1}{2} - (t-2) + \frac{1}{2}(t-2)^2, & t \in [2, 3) \\ 0, & \text{otherwise} \end{cases} \quad (4.4)$$

In figure (Fig. 4.1), we plot the Cardinal B-Splines of degrees 1, 2, 3 and 4.

**R** The colored area under the graph of  $\phi_2$  represents the average  $\int_{x-1}^x \phi_2(t) dt$  which is the value  $\phi_3(x)$ .

### 4.1.1 Cardinal B-Splines properties

Let  $\phi_p$  be a cardinal B-Spline of degree  $p$ ,  $p \in \mathbb{N}$ . The following properties can be proved by induction on the B-Spline degree  $p$ .

**Theorem 4.1.1 — Minimal support.** the support of  $\phi_p$  is  $[0, p + 1]$

**Theorem 4.1.2 — Positivity.**  $\phi_p(s) \geq 0, \forall s \in [0, p + 1]$

**Theorem 4.1.3**  $\phi_p \in \mathcal{C}^{p-1}$

**Theorem 4.1.4**  $\phi_p$  is a piecewise-polynomial of degree  $p$  at each interval  $[i, i + 1], \forall i \in \{0, 1, \dots, p\}$

The sequence  $\{0, 1, 2, \dots, p\}$  is known as the *breaks* of the cardinal B-Spline of degree  $p$ .

**Theorem 4.1.5**  $\forall t \in [0, p + 1]$  and  $p \geq 1$ , we have

$$\dot{\phi}_p(t) = \phi_{p-1}(t) - \phi_{p-1}(t - 1) \quad (4.5)$$

**Theorem 4.1.6 — Symmetry.**  $\phi_p$  is symmetric on the interval  $[0, p + 1]$ , i.e.

$$\phi_p(t) = \phi_p(p + 1 - t), \quad \forall t \in [0, p + 1] \quad (4.6)$$

The following theorem was proved in 1972 by both Cox and Deboor separately.

**Theorem 4.1.7 — Cox-Deboor.**  $\forall t \in [0, p + 1]$  and  $p \geq 1$ , we have

$$\phi_p(t) = \frac{t}{p} \phi_{p-1}(t) + \frac{p+1-t}{p} \phi_{p-1}(t-1) \quad (4.7)$$

*Proof.* Let  $\phi_i^p(t) := \phi_p(t - i), \forall t \in [0, p + 1]$ . We will proof the result by induction. Since both sides vanish at  $t = 0$ , we will use the equivalence to the formula for the derivative 4.1.5.

$$\phi_0^{p-1} - \phi_1^{p-1} = \frac{1}{p} \left( \phi_0^{p-1} - \phi_1^{p-1} \right) + \left[ \frac{t}{p} \left( \phi_0^{p-2} - \phi_1^{p-2} \right) + \frac{p+1-t}{p} \left( \phi_1^{p-2} - \phi_2^{p-2} \right) \right] \quad (4.8)$$

The last term of the previous relation, can be written as

$$\frac{p-1}{p} \left[ \left( \frac{t}{p-1} \phi_0^{p-2} + \frac{p-t}{p-1} \phi_1^{p-2} \right) - \left( \frac{t-1}{p-1} \phi_1^{p-2} + \frac{p-(t-1)}{p-1} \phi_2^{p-2} \right) \right] \quad (4.9)$$

Now, if we assume that the recursion is valid up to  $p - 1$ , then the last terms is equal to

$$\frac{p-1}{p} \left( \phi_0^{p-1} - \phi_1^{p-1} \right) \quad (4.10)$$

■

Interval	Taylor coefficients	
[0, 1]	0 1	$\alpha_{0,1}$
[1, 2]	1 -1	$\alpha_{1,1}$

Table 4.1: The linear Cardinal B-Spline Taylor coefficients.

Interval	Taylor coefficients	
[0, 1]	0 0 $\frac{1}{2}$	$\alpha_{0,2}$
[1, 2]	$\frac{1}{2}$ 1 -1	$\alpha_{1,2}$
[2, 3]	$\frac{1}{2}$ -1 $\frac{1}{2}$	$\alpha_{2,2}$

Table 4.2: The quadratic Cardinal B-Spline Taylor coefficients.

For any Cardinal B-Spline of degree  $p$ , we denote by  $\alpha_i^p = (\alpha_{0,i}^p, \alpha_{1,i}^p, \dots, \alpha_{p,i}^p)$  the sequence of its monomial coefficients on  $[i, i+1]$

$$\phi_p(t) = \alpha_{0,i}^p + \alpha_{1,i}^p t^2 + \dots + \alpha_{p,i}^p t^p, \quad \forall t \in [i, i+1] \quad (4.11)$$

When  $i, j \notin [0, p]$ , we set  $\alpha_{i,j}^p$

**Theorem 4.1.8 — Taylor Coefficients.**

$$\alpha_{i,k}^p = \frac{k}{p} \alpha_{i,k}^{p-1} + \frac{1}{p} \alpha_{i-1,k}^{p-1} + \frac{p+1-k}{p} \alpha_{i,k-1}^{p-1} - \frac{1}{p} \alpha_{i-1,k-1}^{p-1} \quad (4.12)$$

*Proof.* Using the Cox-DeBoor theorem (4.1.7), if  $x = i + t$  where  $t \in [0, 1]$ , we have

$$\begin{aligned} \frac{x}{p} \phi^{p-1}(x) &= \left( \frac{k}{p} + \frac{t}{p} \right) \left( \alpha_{0,i-1}^{p-1} + \alpha_{1,i-1}^{p-1} t^2 + \dots + \alpha_{p,i-1}^{p-1} t^p \right) \\ \frac{p+1-x}{p} \phi^{p-1}(x-1) &= \left( \frac{p+1-k}{p} - \frac{t}{p} \right) \left( \alpha_{0,i-1}^{p-1} + \alpha_{1,i-1}^{p-1} t^2 + \dots + \alpha_{p,i-1}^{p-1} t^p \right) \end{aligned}$$

Adding the last expressions together, we get the the expected relation for  $\alpha_{p,i}^p$ . ■

**Corollary 4.1.9 — Cardinal B-Splines evaluation using pp-form.** Thanks to the theorem (4.1.8), we can compute analytically the Taylor coefficients for low B-Splines order (Tables 4.1, 4.2 and 4.3).

**Theorem 4.1.10 — Inner product.**

$$\int_{\mathbb{R}} \phi_p^{(r)}(t) \phi_q^{(s)}(t+k) dt = (-1)^r \phi_{p+q+1}^{(r+s)}(p+1+k) = (-1)^s \phi_{p+q+1}^{(r+s)}(q+1-k) \quad (4.13)$$

Interval	Taylor coefficients				
[0, 1]	0	0	0	$\frac{1}{6}$	$\alpha_{0,3}$
[1, 2]	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\alpha_{1,3}$
[2, 3]	$\frac{2}{3}$	0	-1	$\frac{1}{2}$	$\alpha_{2,3}$
[3, 4]	$\frac{1}{6}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{6}$	$\alpha_{1,3}$

Table 4.3: The cubic Cardinal B-Spline Taylor coefficients.

## 4.2 Cardinal B-Spline series

### 4.2.1 Scaled and translated Cardinal B-Splines

From now on,  $h > 0$  will denote the mesh step.  $h\mathbb{Z}$  is the uniform grid of width  $h$ . The scaled and translated Cardinal B-Spline of degree  $p$  is defined by

$$\phi_{i,h,p}(x) := \phi_p\left(\frac{x}{h} - i\right) \quad (4.14)$$

The support of  $\phi_{i,h,p}$  is  $[i, i + p + 1]h$ . We introduce the sequence  $(t_i)_{i \in \mathbb{Z}}$ , where  $t_i = ih$ ,  $\forall i \in \mathbb{Z}$ . The following result is another version of the Cox-Deboor theorem (4.1.7).

**Theorem 4.2.1 — Cox-Deboor.**  $\forall t \in \mathbb{R}$  and  $p \geq 1$ , we have

$$\phi_{i,h,p}(t) = \frac{t - t_i}{t_{i+p} - t_i} \phi_{i,h,p-1}(t) + \frac{t_{i+p+1} - t}{t_{i+p+1} - t_{i+1}} \phi_{i+1,h,p-1}(t) \quad (4.15)$$

with  $\phi_{i,h,p}(t) = 0$ ,  $\forall t \notin [t_i, t_{i+p+1}]$ .

*Proof.* Using the definition of the scaled and translated Cardinal B-Spline and the theorem (4.1.7), we have

$$\phi_{i,h,p}(t) = \frac{t - ih}{hp} \phi_{i,h,p-1}(t) + \frac{h(p+1+i) - t}{hp} \phi_{i+1,h,p-1}(t) \quad (4.16)$$

The result is straightforward, using the fact that  $t_{i+p} - t_i = t_{i+p+1} - t_{i+1} = hp$ . ■

### 4.2.2 Cardinal Splines

**Definition 4.2.1 — Cardinal Spline.** A Cardinal Spline, or cardinal B-Spline serie, of degree  $p$  on the grid  $h\mathbb{Z}$  is the linear combination

$$\sum_{k \in \mathbb{Z}} c_k \phi_{k,h,p} \quad (4.17)$$

In figure (Fig. 4.2), we plot all non-vanishing Cardinal B-Splines on the interval  $[2, 3]$ . In figure (Fig. 4.3), we plot a Cardinal B-Spline serie.

**Proposition 4.2.2 — Marsden's identity.**

$$\forall x, t \in \mathbb{R}, (x - t)^p = \sum_{k \in \mathbb{Z}} m_{k,h,p}(t) \phi_{k,h,p}(x) \quad (4.18)$$

where  $m_{k,h,p}(t) = h^p \prod_{i=1}^p \left(k + i - \frac{t}{h}\right)$

*Proof.* ■

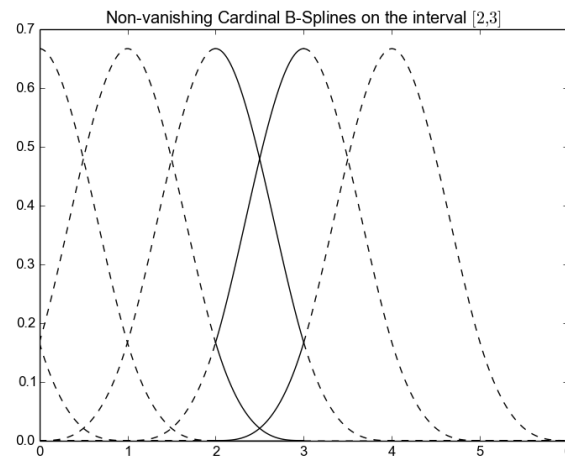
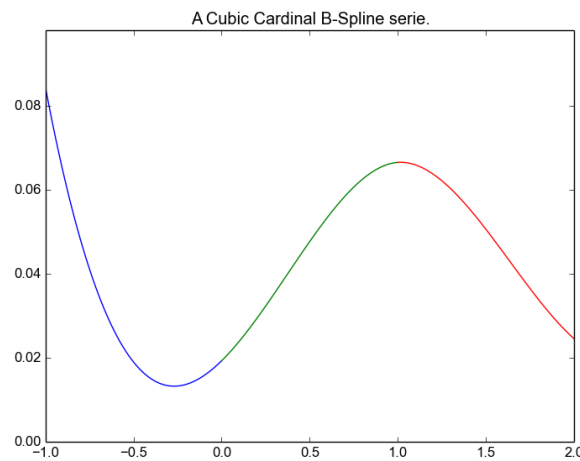
Figure 4.2: Non-vanishing Cardinal B-Splines on the interval  $[2, 3]$ 

Figure 4.3: Example of a cubic Cardinal B-Spline serie.

**Proposition 4.2.3 — Partition of unity.**

$$1 = \sum_{k \in \mathbb{Z}} \phi_{k,h,p}(t), \quad \forall t \in \mathbb{R} \quad (4.19)$$

**Proposition 4.2.4 — Linear independence.** For any element  $[i, i+1]h$ , the Cardinal B-Splines  $(\phi_{k,h,p})_{i-p \leq k \leq i}$  are linearly independent.

**Proposition 4.2.5 — Representation of Polynomials.**

*Proof.* ■

**4.3 Problems**

TODO


## 5. B-Splines curves

### 5.1 B-Splines curves

Let  $(\mathbf{P}_i)_{0 \leq i \leq n} \in \mathbb{R}^d$  be a sequence of control points. Following the same approach as for Bézier curves, we define B-Splines curves as

**Definition 5.1.1 — B-Spline curve.** The B-spline curve in  $\mathbb{R}^d$  associated to  $T = (t_i)_{0 \leq i \leq n+p+1}$  and  $(\mathbf{P}_i)_{0 \leq i \leq n}$  is defined by :

$$\mathcal{C}(t) = \sum_{i=0}^n N_i^p(t) \mathbf{P}_i$$

 The use of open knot vector leads to an **interpolating curve** at the endpoints.



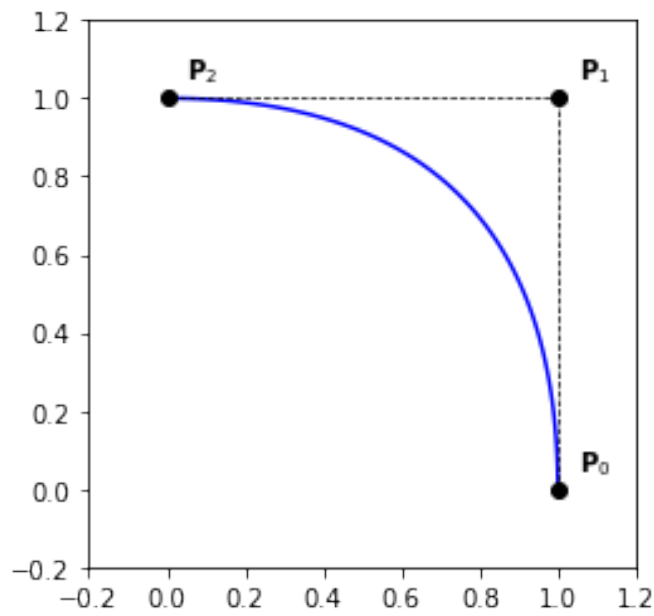


Figure 5.1: Quadratic B-Spline curve using the knot vector  $T = \{0, 0, 0, 1, 1, 1\}$

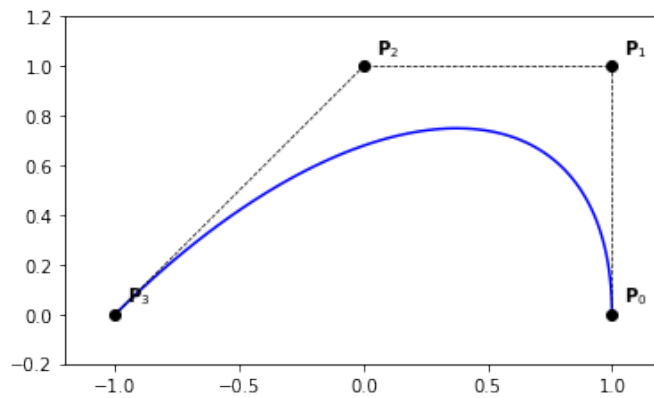


Figure 5.2: Cubic B-Spline curve using the knot vector  $T = \{0, 0, 0, 0, 1, 1, 1, 1\}$

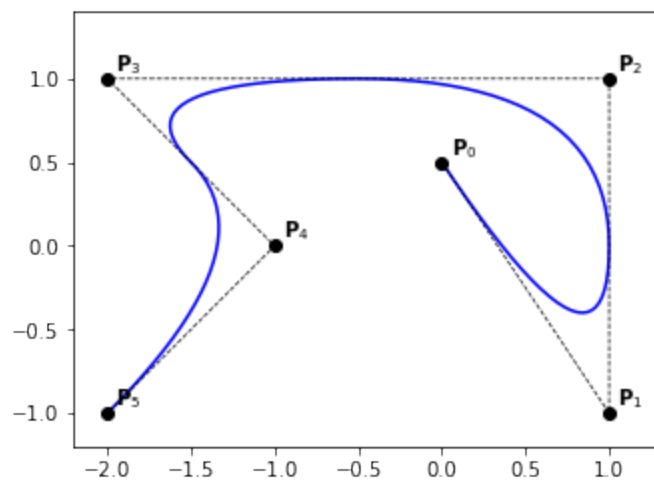


Figure 5.3: Quadratic B-Spline curve using the knot vector  $T = \{0, 0, 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, 1, 1\}$

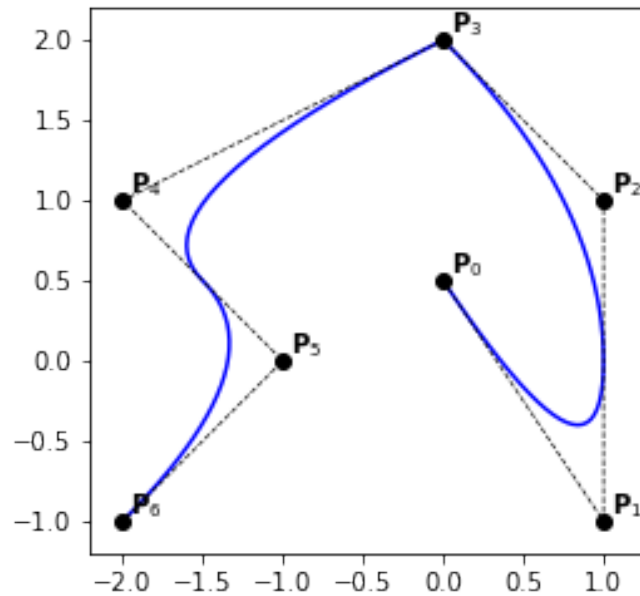


Figure 5.4: Quadratic B-Spline curve using the knot vector  $T = \{0, 0, 0, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}, \frac{3}{4}, 1, 1, 1\}$  with a cusp.

### Examples

**Example 1.** We consider quadratic B-Spline curve on the knot vector  $T = \{0, 0, 0, 1, 1, 1\}$ . This leads to a quadratic Bézier curve.

**Example 2.** We consider cubic B-Spline curve on the knot vector  $T = \{0, 0, 0, 0, 1, 1, 1, 1\}$ . This leads to a cubic Bézier curve.

**Example 3.** We consider quadratic B-Spline curve on the knot vector  $T = \{0, 0, 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, 1, 1\}$ . We remark that

- the curve starts at  $\mathbf{P}_0$  and ends at  $\mathbf{P}_5$ ,
- the curve does not pass through the points  $\{\mathbf{P}_i, 1 \leq i \leq 4\}$ ,
- the tangent directions to the curves at its extremities are parallel to  $\mathbf{P}_1 - \mathbf{P}_0$  and  $\mathbf{P}_5 - \mathbf{P}_4$ ,
- the curve is locally  $\mathcal{C}^2$  and only  $\mathcal{C}^1$  at the knots  $\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$ .

**Example 4.** We consider quadratic B-Spline curve on the knot vector  $T = \{0, 0, 0, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}, \frac{3}{4}, 1, 1, 1\}$ , where the knot  $\frac{1}{2}$  has a multiplicity of 2.

We remark that

- adding a new knot increases the number of control points by 1,
- the curve starts at  $\mathbf{P}_0$  and ends at  $\mathbf{P}_6$  and passes through the point  $\mathbf{P}_3$ ,
- the curve does not pass through the points  $\{\mathbf{P}_i, 1 \leq i \leq 5, i \neq 3\}$ ,
- the tangent directions to the curves at its extremities are parallel to  $\mathbf{P}_1 - \mathbf{P}_0$  and  $\mathbf{P}_6 - \mathbf{P}_5$ ,
- the tangent directions to the curves at  $\mathbf{P}_3$  are parallel to  $\mathbf{P}_3 - \mathbf{P}_2$  and  $\mathbf{P}_4 - \mathbf{P}_3$ ,
- the curve is locally  $\mathcal{C}^2$  inside each subinterval,  $\mathcal{C}^1$  at the knots  $\{\frac{1}{4}, \frac{3}{4}\}$  and only  $\mathcal{C}^0$  at  $\frac{1}{2}$ , where there is a *cusp*.

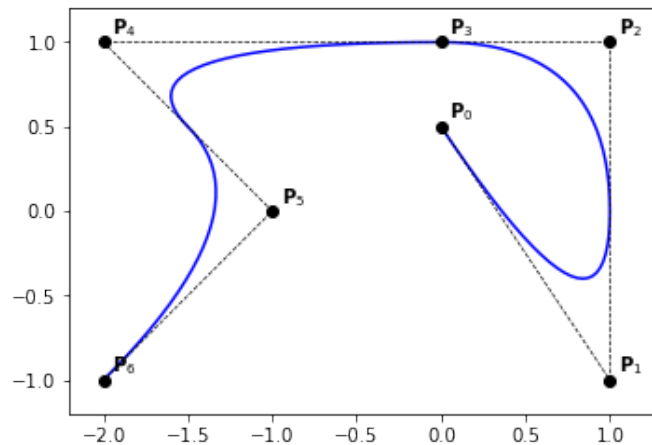


Figure 5.5: Quadratic B-Spline curve using the knot vector  $T = \{0, 0, 0, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}, \frac{3}{4}, 1, 1, 1\}$  without a cusp.

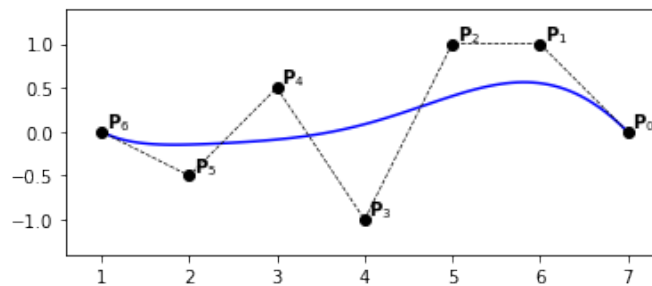


Figure 5.6: B-Spline curve of degree 6 using the knot vector  $T = \{0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1\}$

**Example 5.** We consider quadratic B-Spline curve on the knot vector  $T = \{0, 0, 0, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}, \frac{3}{4}, 1, 1, 1\}$ , where the knot  $\frac{1}{2}$  has a multiplicity of 2.

We remark that although the knot  $\frac{1}{2}$  has a multiplicity of 2, the curve is visually smooth. In fact, it still has the same properties as in the previous example, except that the curve is more than just  $\mathcal{C}^0$  at the double knot  $\frac{1}{2}$ , but not  $\mathcal{C}^1$ . To be more specific, it has a regularity of  $\mathcal{G}^1$  at the knot  $\frac{1}{2}$ .

**Example 6.** In this example, we show the impact of increasing the degree *vs* increasing the number of internal knots.

Given a control polygon defined by the points  $\{P_i, 0 \leq i \leq 6\}$ , we use first a Bézier representation of degree 6 (Fig. 5.6) and a quadratic B-Spline curve (Fig. 5.7) to approach the broken lines defined by the control polygon. As we can see, the use of more knots gives a better approximation. We shall give a proof for this result in the futur.

### Properties of B-Splines curves

We have the following properties for a *B-spline* curve:

- If  $n = p$ , then  $\mathcal{C}$  is a B'ezier-curve,
- $\mathcal{C}$  is a piecewise polynomial curve,
- The curve interpolates its extremas if the associated multiplicity of the first and the last knot are maximum (*i.e.* equal to  $p + 1$ ),

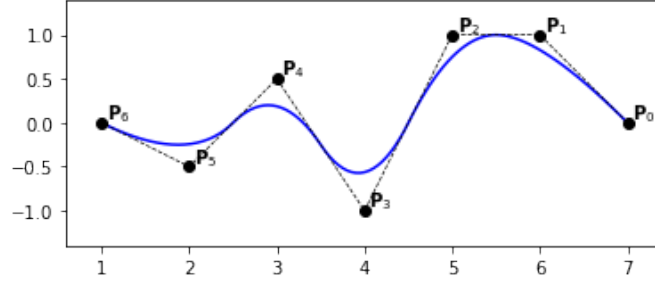


Figure 5.7: Quadratic B-Spline curve using the knot vector  $T = \{0, 0, 0, 0.2, 0.4, 0.6, 0.8, 1, 1, 1\}$

- Invariance with respect to affine transformations,
- B-Spline curves are affinely invariant; *i.e.* the image curve  $\Phi(\sum_{i=0}^n \mathbf{P}_i N_i^p)$  of a B-Spline curve, by an affine mapping  $\Phi$ , is the B-Spline curve having  $(\Phi(\mathbf{P}_i))_{0 \leq i \leq n}$  as control points and the same knot vector,
- strong convex-hull property: if  $t_j \leq t \leq t_{j+1}$ , then  $\mathcal{C}(t)$  is inside the convex-hull associated to the control points  $\mathbf{P}_{j-p}, \dots, \mathbf{P}_j$ ,
- local modification : moving  $\mathbf{P}_j$  affects  $\mathcal{C}(t)$ , only in the interval  $[t_j, t_{j+p+1}]$ ,
- the control polygon is a linear approximation of the curve. We will see later that the control polygon converges to the curve under knot insertion and degree elevation (with different speeds).
- *Variation diminishing*: no plane intersects the curve more than the control polygon.

## 5.2 Derivative of a B-spline curve

Using the derivative formula for *B-spline*, we can compute the derivative of a B-Spline curve:

$$\begin{aligned}
 \mathcal{C}'(t) &= \sum_{i=0}^n \left( N_i^{p'}(t) \mathbf{P}_i \right) \\
 &= \sum_{i=0}^n \left( \frac{p}{t_{i+p} - t_i} N_i^{p-1}(t) \mathbf{P}_i - \frac{p}{t_{i+1+p} - t_{i+1}} N_{i+1}^{p-1}(t) \mathbf{P}_i \right) \\
 &= \sum_{i=0}^n \frac{p}{t_{i+p} - t_i} N_i^{p-1}(t) \mathbf{P}_i - \sum_{i=1}^{n+1} \frac{p}{t_{i+p} - t_i} N_i^{p-1}(t) \mathbf{P}_{i-1} \\
 &= \sum_i N_i^{p-1}(t) \frac{p}{t_{i+p} - t_i} (\mathbf{P}_i - \mathbf{P}_{i-1}) \\
 &= \sum_i N_i^{p-1}(t) \frac{p}{t_{i+p} - t_i} \nabla \mathbf{P}_i
 \end{aligned}$$

where we introduced the **first backward difference** operator  $\nabla \mathbf{P}_i := \mathbf{P}_i - \mathbf{P}_{i-1}$ .

### Examples

**Example 1.** we consider a quadratic B-Spline curve with the knot vector  $T = \{000 \frac{2}{5} \frac{3}{5} 111\}$ . We have  $\mathcal{C}(t) = \sum_{i=0}^4 N_i^2(t) \mathbf{P}_i$ , then

$$\mathcal{C}'(t) = \sum_i N_i^1(t) \mathbf{Q}_i$$

where

$$\mathbf{Q}_0 = 5\{\mathbf{P}_2 - \mathbf{P}_1\} \quad \mathbf{Q}_1 = \frac{10}{3}\{\mathbf{P}_3 - \mathbf{P}_2\}$$

$$\mathbf{Q}_2 = \frac{10}{3}\{\mathbf{P}_4 - \mathbf{P}_3\} \quad \mathbf{Q}_3 = 5\{\mathbf{P}_5 - \mathbf{P}_4\}$$

the *B-splines*  $\{N_i^1, 0 \leq i \leq 3\}$  are associated to the knot vector  $T^* = \{00 \frac{2}{5} \frac{3}{5} 11\}$ .

### 5.3 Rational B-Splines (*NURBS*) curves

Let  $\omega = (\omega_i)_{0 \leq i \leq n}$  be a sequence of non-negative reals. The *NURBS* functions are defined by a projective transformation:

**Definition 5.3.1 — *NURBS* function.** The  $i$ -th *NURBS* of order  $k$ , associated to the knot vector  $T$  and the weights  $\omega$ , is defined by

$$R_i^p = \frac{\omega_i N_i^p}{\sum_j \omega_j N_j^p} \quad (5.1)$$

**R** Notice that when the weights are equal to 1 the *NURBS* are *B-splines*.

**R** *NURBS* functions inherit most of *B-splines* properties. Remark that in the interior of a knot span, all derivatives exist, and are rational functions with non vanishing denominator.

**Definition 5.3.2 — *NURBS* curve.** The *NURBS* curve of degree  $p$  associated to the knot vector  $T$ , the control points  $(\mathbf{P}_i)_{0 \leq i \leq n}$  and the weights  $\omega$ , is defined by

$$\mathcal{C}(t) = \sum_{i=0}^n R_i^p(t) \mathbf{P}_i \quad (5.2)$$

**R** [*NURBS* using perspective mapping] Notice that a *NURBS* curve in  $\mathbb{R}^d$  can be described as a *NURBS* curve in  $\mathbb{R}^{d+1}$  using the control points:

$$\mathbf{P}_i^\omega = (\omega_i \mathbf{P}_i, \omega_i) \quad (5.3)$$

This remark is used for the evaluation and also to extend most of the *B-Splines* fundamental geometric operations to *NURBS* curves.

**R** *NURBS* functions allow us to model, exactly, much more domains than *B-splines*. In fact, all conics can be exactly represented with *NURBS*.

#### 5.3.1 Modeling conics using *NURBS*

In this section, we will show how to construct an arc of conic, using rational *B-splines*. Let us consider the following knot vector :  $T = \{000 \ 111\}$ , the generated *B-splines* are Bernstein polynomials. The general form of a rational Bézier curve of degree 2 is:

$$\mathcal{C}(t) = \frac{\omega_0 N_0^2(t) \mathbf{P}_0 + \omega_1 N_1^2(t) \mathbf{P}_1 + \omega_2 N_2^2(t) \mathbf{P}_2}{\omega_0 N_0^2(t) + \omega_1 N_1^2(t) + \omega_2 N_2^2(t)} \quad (5.4)$$

$\omega = 0$	line
$0 < \omega < 1$	ellipse arc
$\omega = 1$	parabolic arc
$\omega > 1$	hyperbolic arc

Table 5.1: Modeling conics using *NURBS*. Possible values of  $\omega$  and the corresponding conic arc.

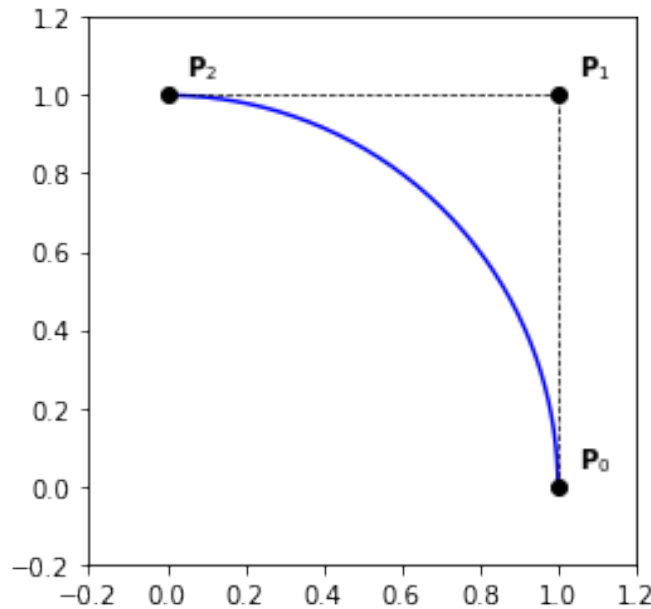


Figure 5.8: Quarter circle using a quadratic Bézier curve.

**R** Because of the multiplicity of the knots 0 and 1, the curve  $\mathcal{C}$  is linking the control point  $\mathbf{P}_0$  to  $\mathbf{P}_2$ .

Let us consider the case  $\omega_1 = \omega_3 = 1$ , in which case the curve will have the following form:

$$\mathcal{C}(t) = \frac{N_0^2(t)\mathbf{P}_0 + \omega N_1^2(t)\mathbf{P}_1 + N_2^2(t)\mathbf{P}_2}{N_0^2(t) + \omega N_1^2(t) + N_2^2(t)} \quad (5.5)$$

Depending on the value of  $\omega$  the resulting Bézier arc is either a line, ellipse, parabolic or hyperbolic arcs (Tab. 5.1).

### Examples

**Example 1.** (Quarter circle) A quarter circle can be described by a quadratic Bézier arc with the control points (Tab. 5.2).

In figure Fig. 5.8, we plot the resulting *NURBS* curve and its control polygon.

**Example 2.** (Circular arc 120) A circular arc of length 120 can be described with a quadratic Bézier curve, using the control points given in Tab. 5.3.

In figure Fig. 5.9, we plot the resulting *NURBS* curve and its control polygon.

**Example 3.** (half circle) In this example we show how to construct half of a circle using a cubic Bézier arc. The control points are given in the Tab. 5.4.

In figure Fig. 5.10, we plot the resulting *NURBS* curve and its control polygon.

	$\mathbf{P}_i$	$\omega_i$
0	(1, 0)	1
1	(1, 1)	$\frac{1}{\sqrt{2}}$
2	(0, 1)	1

Table 5.2: Control points and their associated weights to model a quarter circle quadratic Bézier arc.

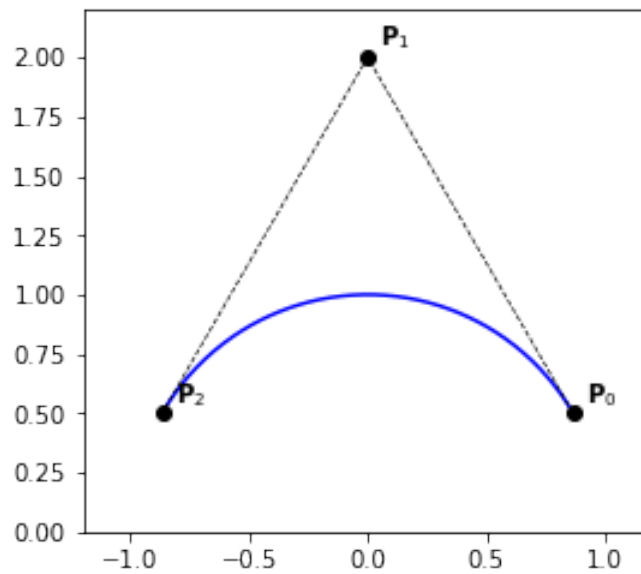


Figure 5.9: A circular arc of length 120 using a quadratic Bézier curve.

	$\mathbf{P}_i$	$\omega_i$
0	$(\cos(\frac{\pi}{6}), \frac{1}{2})$	1
1	(0, 2)	$\frac{1}{2}$
2	$(-\cos(\frac{\pi}{6}), \frac{1}{2})$	1

Table 5.3: Control points and their associated weights to model a circular arc of length 120.

	$\mathbf{P}_i$	$\omega_i$
0	(1, 0)	1
1	(1, 2)	$\frac{1}{3}$
2	(-1, 2)	$\frac{1}{3}$
3	(-1, 0)	1

Table 5.4: Control points and their associated weights to model a half circle using a cubic Bézier arc.

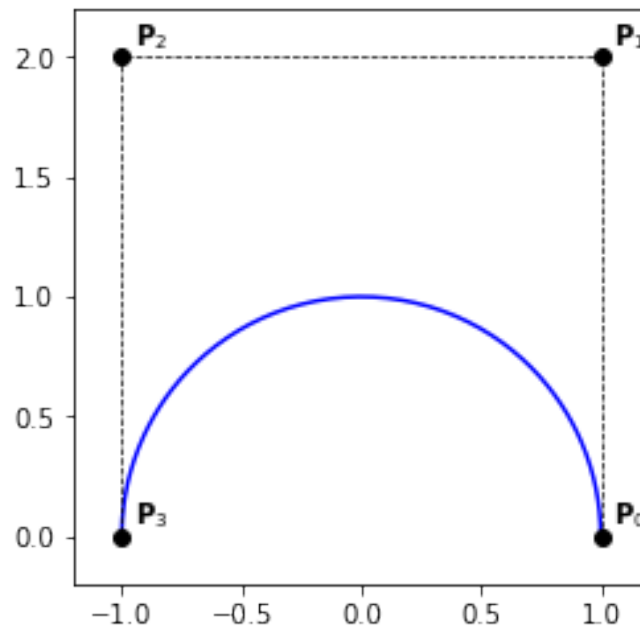


Figure 5.10: Half circle as a cubic Bézier curve.



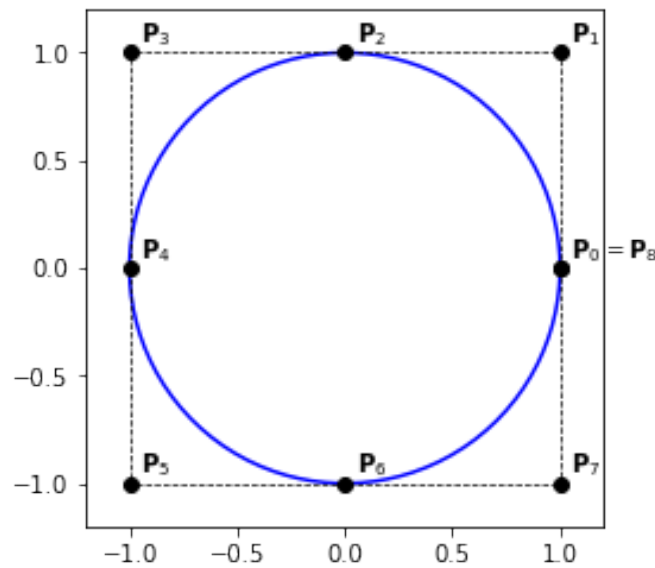


Figure 5.11: Circle as four Bézier curves described by a B-Spline curve using the knot vector  $T = \{0, 0, 0, \frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}, \frac{3}{4}, \frac{3}{4}, 1, 1, 1\}$

	$\mathbf{P}_i$	$\omega_i$
0	(1, 0)	1
1	(1, 1)	$\frac{1}{\sqrt{2}}$
2	(0, 1)	1
3	(-1, 1)	$\frac{1}{\sqrt{2}}$
4	(-1, 0)	1
5	(-1, -1)	$\frac{1}{\sqrt{2}}$
6	(0, -1)	1
7	(1, -1)	$\frac{1}{\sqrt{2}}$
8	(1, 0)	1

Table 5.5: Control points and their associated weights to model a circle using 4 quadratic Bézier arcs.

**Example 4.** (Circle as four arcs) One can draw a circle using four quadratic Bézier arcs, 9 control points (Tab. 5.5) and the knot sequence  $T = \{000, \frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}, \frac{3}{4}, \frac{3}{4}, 111\}$ .

In figure Fig. 5.11, we plot the resulting NURBS curve and its control polygon.

**Example 5.** (Circle as three arcs) One can draw a circle using three quadratic Bézier arcs of length 120, 7 control points (Tab. 5.6) and the knot sequence  $T = \{000, \frac{1}{3}, \frac{1}{3}, \frac{2}{3}, \frac{2}{3}, 111\}$ .

In figure Fig. 5.12, we plot the resulting NURBS curve and its control polygon.

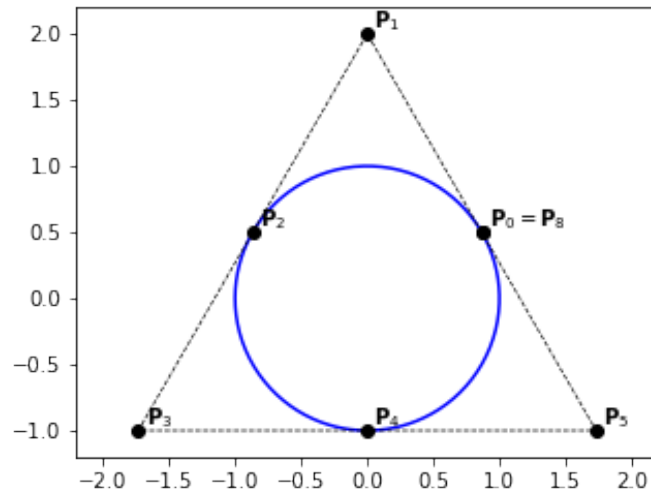


Figure 5.12: Circle as three Bézier curves described by a B-Spline curve using the knot vector  $T = \{000, \frac{1}{3}, \frac{2}{3}, \frac{2}{3}, 111\}$

	$\mathbf{P}_i$	$\omega_i$
0	$(\cos(\frac{\pi}{6}), \frac{1}{2})$	1
1	$(0, 2)$	$\frac{1}{2}$
2	$(-\cos(\frac{\pi}{6}), \frac{1}{2})$	1
3	$(-2\cos(\frac{\pi}{6}), -1)$	$\frac{1}{2}$
4	$(0, -1)$	1
5	$(2\cos(\frac{\pi}{6}), -1)$	$\frac{1}{2}$
6	$(\cos(\frac{\pi}{6}), \frac{1}{2})$	1

Table 5.6: Control points and their associated weights to model a circle using 3 quadratic Bézier arcs of length 1/3.

## 5.4 Fundamental geometric operations

When designing complex CAD models we usually start from simple geometries then manipulate them using different algorithms that are now well established. For example, one may need to have more control on a curve, by adding new control points. This can be done in two different ways:

- insert new knots
- elevate the polynomial degree

In general, one may need to use both approaches.

**R** Notice that these two operations keep the B-Spline curve unchanged.

In the sequel, we will give two algorithms for knot insertion and degree elevation. In fact, there are many algorithms that implement these operations, but we shall only consider one for each operation.

### 5.4.1 Knot insertion

Assuming an initial B-Spline curve defined by:

- its degree  $p$
- knot vector  $T = (t_i)_{0 \leq i \leq n+p+1}$
- control points  $(\mathbf{P}_i)_{0 \leq i \leq n}$

We are interested in the new B-Spline curve as the result of the insertion of the knot  $t$ ,  $m$  times (with a span  $j$ , i.e.  $t_j \leq t < t_{j+1}$ ).

After such operation, the degree is remain unchanged, while the knot vector is enriched by  $t$ ,  $m$  times. The aim is then to compute the new control points  $(\mathbf{Q}_i)_{0 \leq i \leq n+m}$

For this purpose we use the DeBoor algorithm:

$$n := n + m \quad (5.6)$$

$$p := p \quad (5.7)$$

$$T := \{t_0, \dots, t_j, \underbrace{t, \dots, t}_m, t_{j+1}, \dots, t_{n+k}\} \quad (5.8)$$

$$\mathbf{Q}_i := \mathbf{Q}_i^m \quad (5.9)$$

where,

$$\mathbf{Q}_i^0 = \mathbf{P}_i \quad (5.10)$$

$$\mathbf{Q}_i^r = \alpha_i^r \mathbf{Q}_i^{r-1} + (1 - \alpha_i^r) \mathbf{Q}_{i-1}^{r-1} \quad (5.11)$$

with,

$$\alpha_i^r = \begin{cases} 1 & i \leq j - p + r - 1 \\ \frac{t - t_i}{t_{i+p-r+1} - t_i} & j - p + r \leq i \leq j - m \\ 0 & j - m + 1 \leq i \end{cases} \quad (5.12)$$

### Examples

**Example 1.** We consider a cubic B-Spline curve on the knot vector  $T = \{0, 0, 0, 0, 1, 1, 1, 1\}$ .

In this example, we first insert the knot  $\frac{1}{2}$  with a multiplicity of 1. The result is given in Fig. 5.13.

In figures 5.14 and 5.15, we chose a multiplicity of 2 and 3 for the new knot. As a result, we notice

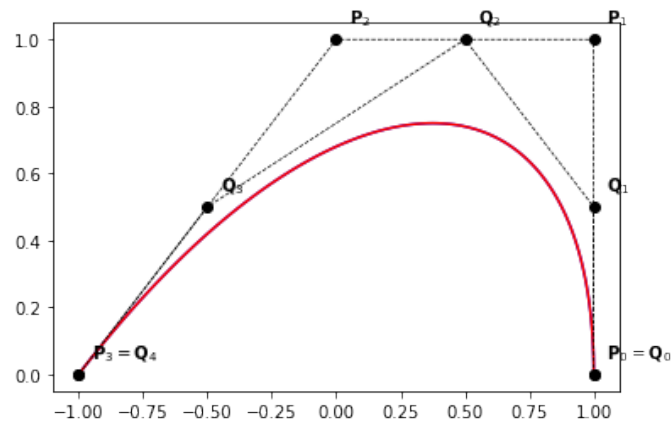


Figure 5.13: A cubic Bézier curve and the new B-Spline curve after inserting the knot  $\frac{1}{2}$  with a multiplicity of 1.

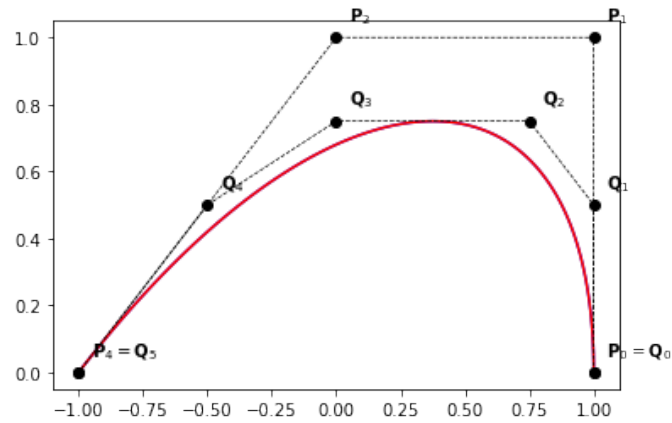


Figure 5.14: A cubic Bézier curve and the new B-Spline curve after inserting the knot  $\frac{1}{2}$  with a multiplicity of 2.

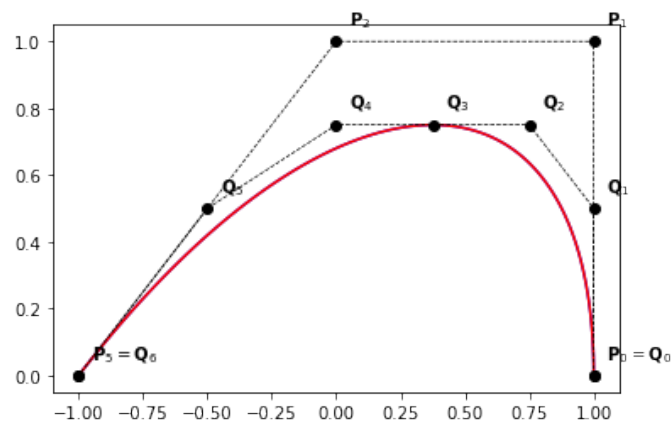


Figure 5.15: A cubic Bézier curve and the new B-Spline curve after inserting the knot  $\frac{1}{2}$  with a multiplicity of 3.

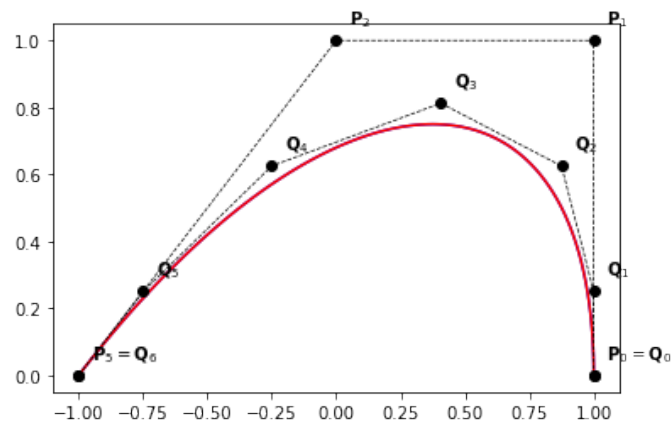


Figure 5.16: A cubic Bézier curve and the new B-Spline curve after inserting the knots  $\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$ .

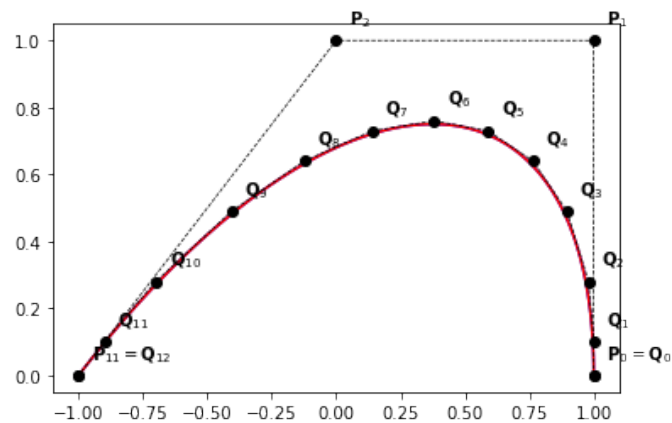


Figure 5.17: A cubic Bézier curve and the new B-Spline curve after inserting the knots  $\{\frac{i}{10}, i \in \{1, 2, \dots, 9\}\}$ .

- whenever we insert a new knot, a new control point appears,
- the curve is unchanged after knot insertion,
- $\mathbf{Q}_1 \in [\mathbf{P}_0\mathbf{P}_1]$  and  $\mathbf{Q}_3 \in [\mathbf{P}_2\mathbf{P}_3]$ .
- in Fig. 5.15, the curve pass through the point  $\mathbf{Q}_3$ , which corresponds to the knot  $\frac{1}{2}$  with a multiplicity of 3. In this case, we subdivided the initial Bézier curve into two Bézier curves.

**Example 2.** We consider again a cubic B-Spline curve on the knot vector  $T = \{0, 0, 0, 0, 1, 1, 1, 1\}$ .

In this example, we show the impact of inserting knots on the control polygon. As we see in figures 5.16 and 5.17, the control polygon converges to the initial curve. We shall prove this result in the Approximation theory part.

For more details about such topic, we refer to [10].

### 5.4.2 Degree elevation

In the sequel, we shall restrict our study to the case of open knot vectors. There are also many algorithms for degree elevation of a B-Spline curve. In the sequel, we will be using the one developed by Huang et al. [13].

Since some knots may be duplicated, we shall assume that the breakpoints are denoted by  $\{t_i^*, 0 \leq i \leq s\}$ , in which case the knot vector has the following form

$$T = \left\{ \underbrace{t_0^*, \dots, t_0^*}_{m_0}, \underbrace{t_1^*, \dots, t_1^*}_{m_1}, \dots, \underbrace{t_{s-1}^*, \dots, t_{s-1}^*}_{m_{s-1}}, \underbrace{t_s^*, \dots, t_s^*}_{m_s} \right\} \quad (5.13)$$

Elevating the degree by  $m$  will lead to the new description (the curve does not change):

$$n := n + sm \quad (5.14)$$

$$p := p + m \quad (5.15)$$

$$T := \left\{ \underbrace{t_0^*, \dots, t_0^*}_{m_0+m}, \underbrace{t_1^*, \dots, t_1^*}_{m_1+m}, \dots, \underbrace{t_{s-1}^*, \dots, t_{s-1}^*}_{m_{s-1}+m}, \underbrace{t_s^*, \dots, t_s^*}_{m_s+m} \right\} \quad (5.16)$$

$$\mathbf{Q}_i := \tilde{\mathbf{Q}}_i^0 \quad (5.17)$$

where the control points  $\tilde{\mathbf{Q}}_i^0$  are given by the following algorithm,

1. we define

$$\beta_i = \sum_{l=1}^i m_l \quad \forall 1 \leq i \leq s-1 \quad (5.18)$$

$$\alpha_i = \prod_{l=1}^i \frac{p-l}{p+m-l} \quad \forall 1 \leq i \leq p-1 \quad (5.19)$$

and set

$$\tilde{\mathbf{P}}_i^0 = \tilde{\mathbf{P}}_i \quad (5.20)$$

2. compute the (scaled) differential coefficients  $\tilde{\mathbf{P}}_i^l$ , for  $l > 0$  as

$$\tilde{\mathbf{P}}_i^l = \begin{cases} \frac{\tilde{\mathbf{P}}_{i+1}^{l-1} - \tilde{\mathbf{P}}_i^{l-1}}{t_{i+p} - t_{i+l}} & , t_{i+p} > t_{i+l} \\ 0 & , t_{i+p} = t_{i+l} \end{cases} \quad (5.21)$$

3. compute for all  $j \in \{0, \dots, p\}$

$$\tilde{\mathbf{P}}_0^j \quad (5.22)$$

4. compute for all  $l \in \{1, \dots, s-1\}$  and  $i \in \{p+1-m_l, \dots, p\}$

$$\tilde{\mathbf{P}}_{\beta_l}^i \quad (5.23)$$

5. compute for all  $j \in \{0, \dots, p\}$

$$\tilde{\mathbf{Q}}_0^j = \prod_{l=1}^j \frac{p+1-l}{p+1+m-l} \tilde{\mathbf{P}}_0^j \quad (5.24)$$

6. compute for all  $l \in \{1, \dots, s-1\}$  and  $i \in \{p+1-m_l, \dots, p\}$

$$\tilde{\mathbf{Q}}_{\beta_l+ml}^j = \prod_{l=1}^j \frac{p+1-l}{p+1+m-l} \tilde{\mathbf{P}}_{\beta_l}^j \quad (5.25)$$

7. compute for all  $l \in \{1, \dots, s-1\}$  and  $i \in \{1, \dots, m\}$

$$\tilde{\mathbf{Q}}_{\beta_l+ml+i}^p = \tilde{\mathbf{Q}}_{\beta_l+ml}^{(p)} \quad (5.26)$$

8. compute  $\tilde{\mathbf{Q}}_i^0$

Note that there exist other algorithms which expand the curve into a Bézier curve, then elevate the degree using Bernstein polynomials, finally come back to a description using *B-splines*. For more details, we refer to [16, 18]. The one given in [13] is more efficient and much more simple to implement. We can also use a more sophisticated version of this algorithm to insert new knots while elevating the degree.

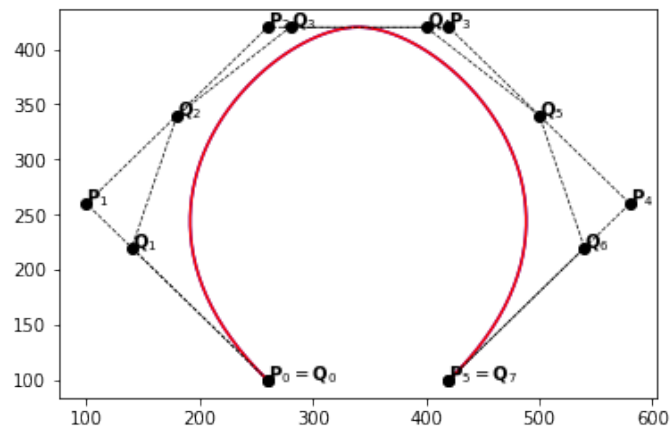


Figure 5.18: A cubic B-Spline curve before and after raising the polynomial degree by 1.

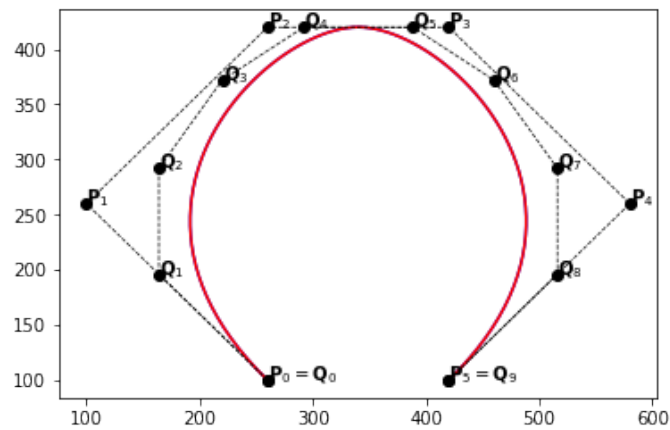


Figure 5.19: A cubic B-Spline curve before and after raising the polynomial degree by 2.

## Examples

**Example 1.** In this example we consider a cubic B-Spline curve with the following knot vector  $T = \{0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 1, 1, 1, 1\}$ .

In Fig. 5.18, we plot the curve before and after raising the B-Spline degree by 1.

We notice that

- in opposition to the knot insertion, elevating the degree by 1 adds two control points rather than 1.
- the curve is unchanged after degree elevation,
- $Q_1 \in [P_0P_1]$  and  $Q_6 \in [P_4P_5]$ .

**Example 2.** In this example we show the convergence of the control polygon under degree elevation. In figures 5.19, 5.20 and 5.21, we plot the control polygon after raising the degree by 2, 4 and 8 respectively.

## 5.5 Problems

TODO



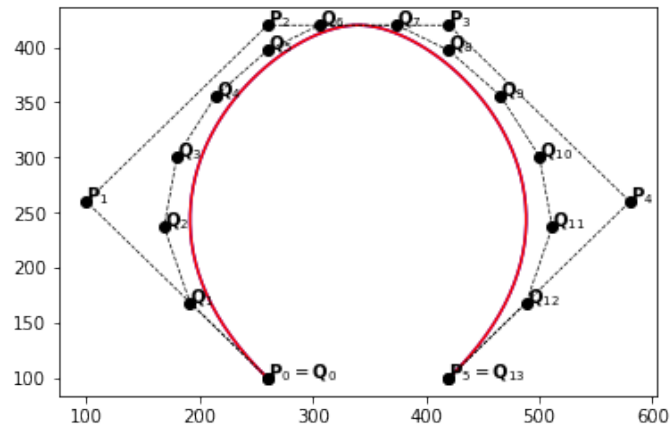


Figure 5.20: A cubic B-Spline curve before and after raising the polynomial degree by 4.

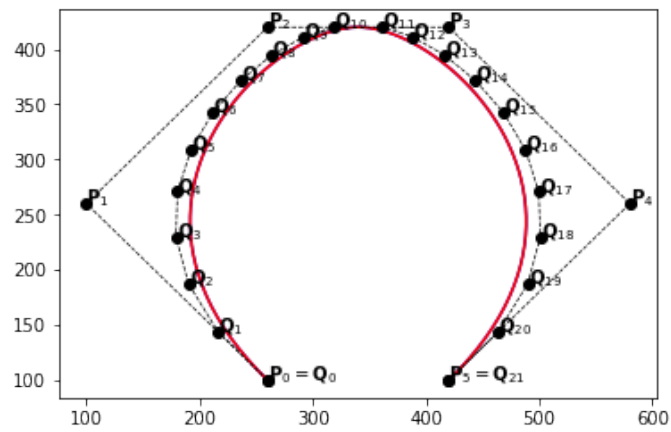


Figure 5.21: A cubic B-Spline curve before and after raising the polynomial degree by 8.

## 6. Historical Notes

References: [16], [3], [4, 5, 6, 17, 19]

Add a general introduction

Cost evaluation of different math functions: ADD, MUL, EXP, SIN, COS, etc



# Approximation theory for B-Splines

<b>7</b>	<b>Divided differences</b> .....	<b>69</b>
7.1	Lagrange interpolation	
7.2	Hermite interpolation	
7.3	Divided differences	
7.4	Problems	
<b>8</b>	<b>Schoenberg space of Spline functions</b>	<b>74</b>
8.1	Basic Splines	
8.2	Spline functions	
8.3	Dual functionals	
8.4	Problems	
<b>9</b>	<b>Spline Approximation</b> .....	<b>79</b>
9.1	Introduction	
9.2	Examples	
9.3	Quasi-Interpolation	
9.4	Global approximation	
9.5	Approximation with Quasi-Interpolation	
9.6	Approximation power of Splines	
9.7	Problems	
<b>10</b>	<b>Historical Notes</b> .....	<b>93</b>



## 7. Divided differences

### 7.1 Lagrange interpolation

Let us consider the interpolation problem for a function  $f$  on a given set of (distinct) points  $\{x_0, \dots, x_n\}$ . It is well known that the Lagrange interpolating polynomial of degree  $n$  writes

$$p(x; x_0, \dots, x_n) = \sum_{i=0}^n f(x_i) L_i(x), \quad \text{where } L_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}, \quad i = 0, \dots, n \quad (7.1)$$

The evaluation of the Lagrange interpolator can be done with different methods. The standard one is known as **Aitken method**. Assume we want to interpolate a function  $f$  on the points  $\{x_0, \dots, x_3\}$ . We start by computing:

$$\begin{aligned} p(x; x_0, x_1) &= \frac{1}{x_1 - x_0} \begin{vmatrix} f(x_0) & x_0 - x \\ f(x_1) & x_1 - x \end{vmatrix} \\ p(x; x_0, x_2) &= \frac{1}{x_2 - x_0} \begin{vmatrix} f(x_0) & x_0 - x \\ f(x_2) & x_2 - x \end{vmatrix} \\ p(x; x_0, x_3) &= \frac{1}{x_3 - x_0} \begin{vmatrix} f(x_0) & x_0 - x \\ f(x_3) & x_3 - x \end{vmatrix} \\ p(x; x_0, x_1, x_3) &= \frac{1}{x_3 - x_1} \begin{vmatrix} p(x; x_0, x_1) & x_1 - x \\ p(x; x_0, x_3) & x_3 - x \end{vmatrix} \\ p(x; x_0, x_1, x_2) &= \frac{1}{x_2 - x_1} \begin{vmatrix} p(x; x_0, x_1) & x_1 - x \\ p(x; x_0, x_2) & x_2 - x \end{vmatrix} \end{aligned}$$

Then the value of the interpolation polynomial of degree 3 at  $x$  is given by

$$p(x; x_0, x_1, x_2, x_3) = \frac{1}{x_3 - x_2} \begin{vmatrix} p(x; x_0, x_1, x_2) & x_2 - x \\ p(x; x_0, x_1, x_3) & x_3 - x \end{vmatrix}$$

The complexity of the Aitken method is  $O(n^2)$  which is expensive, since we know that the Horner algorithm is only about  $O(n)$ .

### 7.1.1 Newton form and Neville's algorithm

Another interesting form is the so called **Newtonian form** of  $p$  and writes

$$p(x) = \sum_{i=0}^n a_i \prod_{j=0}^{i-1} (x - x_j) \quad (7.2)$$

Let us explain how to use the form 7.2 to have a fast evaluation of the interpolant polynomial at a given point  $x$ .

First, we define the polynomial  $p_{j,k}$  as the interpolant polynomials of degree  $k$  at the sites  $\{x_j, x_{j+1}, \dots, x_{j+k}\}$ , i.e.  $p_{j,k}(x_i) = f(x_i)$  for  $i \in \{j, j+1, \dots, j+k\}$ . Where we assume the sites to be ordered and distincts. Therefor, these polynomials exist and are unique.

In the linear case, we have

$$p_{0,1}(x) = f(x_0) + (x - x_0) \frac{f(x_1) - f(x_0)}{x_1 - x_0} = a_0 + (x - x_0)a_1 \quad (7.3)$$

for some coefficients  $a_0$  and  $a_1$ . In this case, we have

$$a_0 = f(x_0) \quad (7.4)$$

and

$$a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} \quad (7.5)$$

$a_1$  is called the **first order divided difference** of  $\mathcal{X} := \{f(x_0), \dots, f(x_n)\}$ . In the quadratic case, we can write

$$p_{0,2}(x) = p_{0,1}(x) + (x - x_0)(x - x_1)a_2 \quad (7.6)$$

where

$$a_2 = \frac{f(x_2) - p_{0,1}(x_2)}{(x_2 - x_0)(x_2 - x_1)} \quad (7.7)$$

$a_2$  is called the **second order divided difference** of  $\mathcal{X}$ . We notice that, we can write, for a given coefficient  $a$ ,

$$p_{j,k}(x) = p_{j,k-1}(x) + (x - x_j) \dots (x - x_{j+k-1})a$$

The coefficient  $a$  is the  **$k$ -th order divided difference** of  $\mathcal{X}$  and will be denoted  $[x_j, \dots, x_{j+k}]f$ . Since,

$$p_{j,k}(x) = \frac{x_{j+k} - x}{x_{j+k} - x_j} p_{j-1,k-1}(x) + \frac{x - x_j}{x_{j+k} - x_j} p_{j,k-1}(x), \quad j \in \{0, \dots, p-k\}$$

we get the following result, by comparing the coefficients of the monomial  $x^k$ ,

$$[x_j, \dots, x_{j+k}]f := \frac{1}{x_{j+k} - x_j} ([x_{j+1}, \dots, x_{j+k}]f - [x_j, \dots, x_{j+k-1}]f) \quad (7.8)$$

The Neville's algorithm is then given as follows

1. set  $[x_j]f := f(x_j)$ , for all  $j \in \{0, \dots, p\}$ ,
2. use 7.8 to compute  $[x_j, \dots, x_{j+k}]f$ , for all  $j \in \{0, \dots, p-k\}$  and  $k \in \{1, \dots, p\}$

### Horner algorithm

We give here a modified version for the Horner algorithm based on the Newtonian form.

1. set  $q := a_p$
2. for every  $k \in \{p-1, p-2, \dots, 0\}$ , we update  $q$  using  $q := a_k + (x - x_k)q$
3. the evaluation of the polynomial at  $x$  is then given by  $p_{0,p}(x) := q$

A quick study of the complexity leads to  $p$  multiplications and  $2p$  additions and subtractions.

## 7.2 Hermite interpolation

In the Hermite interpolation, not only the values of a function are given, but also some of its successive derivatives. Given the set of interpolation points  $X := \{x_0, \dots, x_n\}$ , we construct the set of distinct points  $X^* := \{x_0^*, \dots, x_s^*\}$ , for which we associate a multiplicity  $m_j > 0$  with  $\sum_{j=0}^s m_j = n + 1$ . Let  $c_{j,l}$  be given constants. The Hermite interpolation problem is

**Definition 7.2.1 — Hermite interpolation.** Find a polynomial  $H_n \in \Pi_n$  such that

$$H_n^{(l)}(x_j^*) = c_{j,l} \quad 0 \leq l \leq m_j - 1 \quad \text{and} \quad 0 \leq j \leq s \quad (7.9)$$

It is easy to prove that there exist a unique polynomial  $H_n$  that solves the Hermite interpolation problem.

In the sequel, we shall define the set of constraints by reordering the coefficients  $c_{j,l}$  such that  $(d_j)_{j=0}^n := \{c_{j,l}, 0 \leq j \leq s, 0 \leq l \leq m_j - 1\}$ . Now we can state the general theorem for the Newton's method:

**Theorem 7.2.1 — Newton's method.** There exist unique constants  $a_0, \dots, a_n$  for which the polynomials

$$\begin{cases} P_0(x) = a_0 \\ P_1(x) = a_0 + a_1(x - x_0) \\ P_2(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) \\ \dots \\ P_n(x) = a_0 + a_1(x - x_0) + \dots + a_n \prod_{j=0}^{n-1} (x - x_j) \end{cases} \quad (7.10)$$

are solutions of the Hermite interpolation problems for the sets of points  $\{x_0\}$ ,  $\{x_0, x_1\}$ ,  $\dots$ ,  $\{x_0, \dots, x_n\}$  and given data  $(d_j)_{j=0}^n$ .

**R** The Hermite interpolation is a generalization of both Lagrange interpolation and the Taylor interpolation. Recall that Taylor interpolant of a smooth function  $f \in \mathcal{C}^n([a, b])$  is defined by

$$T_n(x) := f(x_0) + f'(x_0)(x - x_0) + \dots + f^{(n)}(x_0) \frac{(x - x_0)^n}{n!} \quad (7.11)$$

These are the two extreme cases of Hermite interpolation where for the Lagrange interpolation the interpolation points are all distinct, while for Taylor interpolation there all equal.

## 7.3 Divided differences

**Definition 7.3.1 — Divided Differences (G. Kowalewski 1932).** For a set of points (not necessarily ordered)  $X := \{x_0, \dots, x_n\}$ , and a function  $f$ , we define the  $n$ -th divided difference of  $f$  by

$$[x_0, \dots, x_n]f := a_n \quad (7.12)$$

where  $a_n$  is the coefficient of  $x^n$  of the polynomial which interpolates  $f$  at  $x_0, \dots, x_n$  as shown in the Newtonian form 7.2.

Now let's go back to the divided differences operator. After giving additional examples, we shall present some properties.

### Examples

**Example 1.**  $[x_0]f = f(x_0)$

**Example 2.**  $[x_0, x_1]f = \frac{f(x_0) - f(x_1)}{x_0 - x_1}$  if  $x_0 \neq x_1$

**Example 3.**  $[x_0, x_0]f = f'(x_0)$

Because of the symmetry of the Newton form we have

**Proposition 7.3.1**  $[x_0, \dots, x_n]f$  is symmetric in  $x_0, \dots, x_n$ .

From the Newton form, we also deduce

**Proposition 7.3.2**  $[x_0, \dots, x_n]f$  is constant if  $f$  is a polynomial of degree  $\leq n$ , and zero for a polynomial of degree  $< n$ .

Use the Taylor polynomial we have

**Proposition 7.3.3**  $[x_0, \dots, x_0]f = \frac{1}{n!}f^{(n)}(x_0)$ .

**Proposition 7.3.4**  $[x_0, \dots, x_n]f$  is a linear combination of the derivatives  $f^{(l)}(x_i)$ ,  $0 \leq l \leq m_i - 1$ , where  $m_i$  is the multiplicity of the point  $x_i$  in the set  $X$ .

**Proposition 7.3.5** if  $f \in \mathcal{C}^n([a, b])$ ,  $a \leq x_i \leq b$ ,  $0 \leq i \leq n$ , then :

$$[x_0, \dots, x_n]f = \frac{1}{n!}f^{(n)}(\xi), \quad \text{for some } \xi \in [a, b]$$

**Proposition 7.3.6**  $[x_0, \dots, x_n]f$  is continuous at the sites in  $X$ , if the derivatives of  $f$  of proper orders are continuous at the considered site.

**Proposition 7.3.7** if  $x_0 \neq x_n$ , we have

$$[x_0, \dots, x_n]f = \frac{1}{x_n - x_0} \{ [x_1, \dots, x_n]f - [x_0, \dots, x_{n-1}]f \} \quad (7.13)$$

**Proposition 7.3.8 — Leibniz's Formula.**

$$[x_0, \dots, x_n](fg) = \sum_{i=0}^n [x_0, \dots, x_i](f) [x_i, \dots, x_n](g)$$

**Proposition 7.3.9** If  $f^{(n-1)}$  is absolutely continuous, and if not all  $x_i$  coincide, we have

$$[x_0, \dots, x_n]f = \int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} f^{(n)}(x_0 + h_1 t_1 + h_2 t_2 + \cdots + h_n t_n) dt_n$$

where we denote  $h_i = x_{i+1} - x_i$ ,  $i \in \{0, \dots, n-1\}$ .



**Corollary 7.3.10**

$$|[x_0, \dots, x_n]f| \leq \frac{1}{n!} \|f^{(n)}\|_\infty$$

This shows, that the functional  $f \rightarrow [x_0, \dots, x_n]f$  is continuous on  $\mathcal{C}^n[a, b]$ .

**Proposition 7.3.11** If all  $x_i$  are distinct, then

$$[x_0, \dots, x_n]f = \int_a^b f^{(n)}(t)[x_0, \dots, x_n] \left( \frac{(\cdot - t)_+^{n-1}}{(n-1)!} \right) dt$$

This gives a representation of the functional  $[x_0, \dots, x_n]$  in term of the Peano kernel.

**7.4 Problems**

TODO

## 8. Schoenberg space of Spline functions

### 8.1 Basic Splines

In the previous chapters, we introduced the B-Splines through the Cox-DeBoor formula. In this section, we shall give a more specific definition, using the Divided Differences operator.

**Definition 8.1.1 — M-Spline.** Let  $X = \{x_0, \dots, x_{p+1}\}$  a non-decreasing sequence of  $p+2$  points such that  $x_0 \neq x_{p+1}$ . The M-Spline will be defined in term of the following Divided-Difference :

$$M(x) = M(x; X) := M(x; x_0, \dots, x_{p+1}) = (p+1)[x_0, \dots, x_{p+1}](\cdot - x)_+^p \quad (8.1)$$

The points forming the set  $X$ , are called *knots*, and  $X$  is said to be a *knot vector*.

**Proposition 8.1.1**  $M(x) = 0, \forall x \notin [x_0, x_{p+1}]$

**Proposition 8.1.2**  $M(x) > 0, \forall x \in (x_0, x_{p+1})$

**Proposition 8.1.3**  $\frac{M}{n!}$  is the Peano kernel of the divided-difference at the set  $X = \{x_0, \dots, x_{p+1}\}$ . Then, for any  $f \in W_1^{p+1}$ , we have,

$$[x_0, \dots, x_{p+1}]f = \int_{-\infty}^{+\infty} f^{(p+1)}(t)M(t)dt$$

**Proposition 8.1.4**  $\int_{-\infty}^{+\infty} M(t)dt = 1$

**Proposition 8.1.5** If  $p \geq 1$ , we have the recurrence formula,

$$M(x; x_0, \dots, x_{p+1}) = \frac{p+1}{p} \left( \frac{x-x_0}{x_{p+1}-x_0} M(x; x_0, \dots, x_p) + \frac{x_{p+1}-x}{x_{p+1}-x_0} M(x; x_1, \dots, x_{p+1}) \right) \quad (8.2)$$

The B-Splines are then defined in terms of M-Splines as

$$N(x; x_0, \dots, x_{p+1}) = \frac{x_{p+1}-x_0}{p+1} M(x; x_0, \dots, x_{p+1}) \quad (8.3)$$

In this case we recover the Cox-DeBoor formula

$$N(x; x_0, \dots, x_{p+1}) = \frac{x-x_0}{x_p-x_0} N(x; x_0, \dots, x_p) + \frac{x_{p+1}-x}{x_{p+1}-x_1} N(x; x_1, \dots, x_{p+1}) \quad (8.4)$$

In fact, one can see the property 8.1.4 of the M-Splines as a normalization of the basic splines, such that the integral is one. Another normalization is to have the partition unity, which is given by the B-Splines.

### 8.1.1 Smoothness of a B-Spline

**Proposition 8.1.6** Suppose that  $x^* \in X := \{x_0, \dots, x_{p+1}\}$  occurs  $m$  times in the knot sequence in  $X$ . If  $1 \leq m \leq p+1$ , then  $N^{(r)}$  is continuous at  $x^*$  for all  $r \in \{0, \dots, p-m\}$ , while  $N^{p-m+1}$  is discontinuous at  $x^*$ .

*Proof.* TODO ■

## 8.2 Spline functions

Splines are piecewise polynomials defined on the real line. We shall require that on each compact interval, they consist of a small number of non-vanishing polynomial pieces.

Let  $T^* = \{t_i^*, 0 \leq i \leq s\}$  be a finite strictly increasing sequence of points of  $\mathbb{R}$ . A function  $S$  on  $\mathbb{R}$  is a **spline of degree  $p, p \geq 0$**  with the breakpoints  $T^*$  if on each interval  $(t_i^*, t_{i+1}^*)$ , it is a **polynomial of degree  $\leq p$** . At every breakpoint  $t_i^*$ ,  $S$  and its derivatives (which are also splines) are defined by continuity whenever it is possible. For example, splines of order one are step functions, those of order two are broken lines.

At every breakpoint  $t_i^*$ , a spline  $S$  shall have a **smoothness  $r_i$** , which is defined as the following:

- $r_i := -1$  if  $S$  is discontinuous at  $t_i^*$ , otherwise,
- $r_i$  is such that  $S^{(j)}$  is well defined and continuous at  $t_i^*$  for all  $0 \leq j \leq r_i$  and  $1 \leq i \leq s-1$ .

**R** We will not consider the case of discontinuous splines at internal breakpoints. Meaning that  $r_i \geq 0$ .

**R**  $k_i := p - r_i$  is called the **defect** and is the number of degrees of freedom of  $S$  at  $t_i^*$

In the sequel, we shall consider the case of the interval  $I = [a, b]$ , where  $a := t_0^*$  and  $b = t_s^*$ .

**Definition 8.2.1 — Schoenberg space.** Given a set of breakpoints  $T^*$  as defined previously and a sequence of integers  $\mathbf{r} := \{r_i, 1 \leq i \leq s-1\}$ . The space of functions which consists of all splines  $S$  of degree  $\leq p$  with breakpoints contained in  $T^*$  and of smoothness  $\geq r_i$  at  $t_i^*$  is called **Schoenberg space**, which we denote  $\mathcal{S}_p^{\mathbf{r}}(T^*)$ .

An example of a spline function is the *truncated power*, if  $i > 0$  and  $\tau \in T^*$ ,

$$(x - \tau)_+^i := \begin{cases} (x - \tau)^i, & x \geq \tau \\ 0, & x < \tau \end{cases} \quad (8.5)$$

**Exercise 8.1** Show that the truncated powers  $(x - t_i^*)_+^j \in \mathcal{S}_p^{\mathbf{r}}(T^*)$ , for all  $j \in \{r_i, \dots, p\}$ . ■

Using the truncated powers, one can construct a basis for the Schoenberg space. However, this basis will not have the compact support property. We shall see in the next sections, that the B-Splines form a basis for the Schoenberg space. There are different methods to prove this fact, the one that we will follow is based on the DeBoor-Fix quasi-interpolation approach.

### 8.3 Dual functionals

In this section, we are interested in the construction of dual functionals  $(\lambda_i)_{0 \leq i \leq n}$  for a family of B-Splines  $(N_j^p)_{0 \leq j \leq n}$  such that

$$\lambda_i(N_j^p) = \delta_{ij}, \quad \forall i, j \in \{0, \dots, n\} \quad (8.6)$$

We recall the Mardsen's identity 3.3:

$$(x-y)^p = \sum_j \psi_{j,p}(y) N_j^p(x)$$

where

$$\psi_{j,p}(y) = \prod_{i=j+1}^{j+p} (t_i - y) \quad (8.7)$$

**Proposition 8.3.1 — representation of polynomials.** Let  $g \in \Pi_p$ , then

$$g(x) = \sum_{j=0}^n \mu_j^p(g) N_j^p(x) \quad (8.8)$$

where

$$\mu_j^p(g) = \frac{1}{p!} \sum_{r=0}^p (-1)^{p-r} \psi_{j,p}^{(r)}(\tau_j) g^{(p-r)}(\tau_j) \quad (8.9)$$

for some  $\tau_j \in [a, b]$ .

*Proof.* Let  $g \in \Pi_p$  and  $\tau_j \in [a, b]$ . Using the Taylor expansion (or Taylor form of the polynomial) we get

$$g(x) = \sum_{r=0}^p \frac{1}{(p-r)!} (x - \tau_j)^{p-r} g^{(p-r)}(\tau_j) \quad (8.10)$$

but using the Mardsen identity for  $y = \tau_j$ , we have

$$(x - \tau_j)^{p-r} = \sum_{j=0}^n (-1)^{p-r} \frac{(p-r)!}{p!} \psi_{j,p}^{(r)}(\tau_j) N_j^p(x)$$

by substituting this expression in Eq. 8.10, we get

$$g(x) = \sum_{r=0}^p \frac{1}{(p-r)!} \sum_{j=0}^n (-1)^{p-r} \frac{(p-r)!}{p!} \psi_{j,p}^{(r)}(\tau_j) N_j^p(x) g^{(p-r)}(\tau_j)$$

which leads to

$$g(x) = \sum_{j=0}^n \frac{1}{p!} \sum_{r=0}^p (-1)^{p-r} \psi_{j,p}^{(r)}(\tau_j) g^{(p-r)}(\tau_j) N_j^p(x)$$

■

Next, we give a general representation of any spline function in terms of B-Splines.

**Proposition 8.3.2 — representation of splines.** Let  $g \in \mathcal{S}_p(T)$ , then

$$g(x) = \sum_{j=0}^n \lambda_j^p(g) N_j^p(x) \quad (8.11)$$

where

$$\lambda_j^p(g) = \frac{1}{p!} \begin{cases} \sum_{r=0}^p (-1)^{p-r} \psi_{j,p}^{(r)}(\tau_j) D_+^{p-r} g(\tau_j) & , \tau_j = t_j \\ \sum_{r=0}^p (-1)^{p-r} \psi_{j,p}^{(r)}(\tau_j) g^{(p-r)}(\tau_j) & , t_j < \tau_j < t_{j+p+1} \\ \sum_{r=0}^p (-1)^{p-r} \psi_{j,p}^{(r)}(\tau_j) D_-^{p-r} g(\tau_j) & , \tau_j = t_{j+p+1} \end{cases} \quad (8.12)$$

*Proof.* Let  $\tau_j \in [t_j, t_{j+p+1}]$ . We also denote by  $m_j$  the integer such that  $\tau_j \in I_j := [t_{m_j}, t_{m_j+1}]$ . There are three cases to consider: either  $\tau_j$  is inside the interval  $I_j$  or is equal to one of its bounds. For the moment, let us consider the case where  $t_{m_j} < \tau_j < t_{m_j+1}$ .

Since  $s|_{I_j} \in \Pi_p$ , we use the representation of polynomials to get

$$s|_{I_j} = \sum_{i=m_j-p}^{m_j} \mu_i^p(s|_{I_j}) N_i^p$$

on the other hand, since  $j \leq m_j \leq j+p$ , we have  $m_j - p \leq j \leq m_j$  and using the local independence of B-Splines on  $I_j$ , we get

$$\lambda_j^p(s) = \mu_j^p(s|_{I_j})$$

The other cases are treated in the same way by replacing the derivatives on  $\tau_j$  by the right (left) derivatives if  $\tau_j = t_j$  (or  $\tau_j = t_{j+p+1}$ ). ■

**R** If we denote by  $\mu_j$  the multiplicity of  $\tau_j$  in  $\{t_{j+1}, \dots, t_{j+p}\}$ , then  $\psi_{j,p}^{(r)}(\tau_j) = 0$  for all  $r \leq \mu_j - 1$ . Therefore, the summation index in 8.12 starts from  $\mu_j$  rather than 0.

Let  $T$  be a knot sequence associated to a set of breakpoints  $T^*$  for which each internal knot has a multiplicity  $m_i$ . From Prop 8.1.6, we know that a B-Spline  $N_j$  has  $p - m_i$  continuous derivatives at a breakpoint  $t_i^*$  with a multiplicity  $m_i$ , which lies in the support of  $N_j$ .

**Proposition 8.3.3 — B-Splines as a basis for Schoenberg space.** TODO

*Proof.* TODO ■

## 8.4 Problems

TODO

# 9. Spline Approximation

## 9.1 Introduction

The aim of this chapter is to construct a spline approximation of a function  $f$ , in terms of B-Splines as

$$\mathcal{L}f(x) = \sum_{j=0}^n \lambda_j(f) N_j^p(x) \quad (9.1)$$

Depending on the nature of the functionals  $\lambda_i$  the approximation will be global (interpolation and least square approximation) or local (quasi-interpolation).

## 9.2 Examples

### 9.2.1 Piecewise linear interpolation

We consider in this example, a set of increasing points (which will represent our grid points)  $x_0 < x_1 < \dots < x_m$ . We also denote  $a = x_0$  and  $b = x_m$ . We first construct a knot vector from the grid points, by duplicating the first and the last points

$$T_1 := \{x_0, x_0, x_1, \dots, x_{m-1}, x_m, x_m\}$$

We consider the set of linear B-Splines associated to  $T_1$ , denoted by  $(N_j^1)_{j=0}^m$ .

**Proposition 9.2.1** Let  $f \in \mathcal{C}[a, b]$ . The spline defined as

$$I_1 f(x) = \sum_{j=0}^m f(x_j) N_j^1(x) \quad (9.2)$$

satisfies the interpolation condition, *i.e.*

$$I_1 f(x_i) = f(x_i) \quad \forall i \in \{0, \dots, m\}$$

**Exercise 9.1** Prove the previous proposition. ■

### 9.2.2 Variation diminishing spline approximation

In the sequel, we present an approximation that preserves the bounds, monotonicity and convexity. In addition, it reproduces all linear piecewise polynomials.

**Definition 9.2.1 — Variation diminishing approximation.** Let  $f \in \mathcal{C}[a, b]$  and  $p$  a positive integer. Moreover, we consider a  $p + 1$ -regular knot sequence  $T = \{t_0, \dots, t_{n+p+1}\}$ , with  $t_p = a$  and  $t_{n+1} = b$ . We also define the knot averages, also known as the Greville points,

$$t_j^* = \frac{1}{p} (t_{j+1} + \dots + t_{j+p}) \quad (9.3)$$

The Variation diminishing spline approximation of degree  $p$  to  $f$ , on the knot sequence  $T$ , is defined as

$$Vf(x) = \sum_{j=0}^n f(t_j^*) N_j^p(x) \quad (9.4)$$

**R** The first knot  $t_0$  and the last one  $t_{n+p+1}$  do not appear in the definition of the Greville points.

**R** If all internal knots occur less than  $p + 1$ , i.e. the Schoenberg space  $\mathcal{S}_p(T) \subset \mathcal{C}[a, b]$ , then

$$t_0^* < t_1^* < \dots < t_n^*$$

#### Preserving the bounds

**Lemma 2** If  $g = \sum_{j=0}^n c_j N_j^p$  is an element of the Schoenberg space  $\mathcal{S}_p(T)$ , then

$$\min_{0 \leq j \leq n} c_j \leq g \leq \max_{0 \leq j \leq n} c_j$$

*Proof.* Since the B-Splines are positive, we have

$$\min_{0 \leq j \leq n} c_j \left( \sum_{j=0}^n N_j^p \right) \leq g \leq \max_{0 \leq j \leq n} c_j \left( \sum_{j=0}^n N_j^p \right)$$

we conclude by using the fact that the B-Splines forms a partition of unity. ■

**Proposition 9.2.2** Let  $f \in \mathcal{C}[a, b]$  such that  $m \leq \|f\|_\infty \leq M$  for some real numbers  $m$  and  $M$ . Then

$$m \leq \|Vf\|_\infty \leq M \quad (9.5)$$

*Proof.* Since  $Vf(x) = \sum_{j=0}^n f(t_j^*) N_j^p(x)$  and for all  $0 \leq j \leq n$  we have  $m \leq f(t_j^*) \leq M$ . We conclude by using the previous lemma. ■



### Preserving the monotonicity

**Lemma 3** If  $g = \sum_{j=0}^n c_j N_j^p$  is an element of the Schoenberg space  $\mathcal{S}_p(T)$ , then

1. if  $\forall j \ 0 \leq j \leq n-1, c_j \leq c_{j+1}$  then  $g$  is increasing
2. if  $\forall j \ 0 \leq j \leq n-1, c_j \geq c_{j+1}$  then  $g$  is decreasing

*Proof.* We shall assume  $p \geq 1$ , since the case  $p = 0$  is straightforward. We shall also give the proof only for the first point, the second case is treated in a similar way. Let  $x \in [a, b]$ , we have

$$g'(x) = \sum_j \nabla c_j N_j^{p-1}(x)$$

with  $\nabla c_j := \frac{c}{t_{j+p} - t_j} (c_j - c_{j-1})$ .

But since  $\nabla c_j \geq 0$ , then  $g'(x) \geq 0$ . Therefore  $g$  is increasing. ■

**Proposition 9.2.3** If  $f \in \mathcal{C}[a, b]$  is increasing (decreasing) then  $Vf$  is increasing (decreasing).

*Proof.* We use the fact that the coefficients in the Variation diminishing approximation are the evaluation of the function  $f$  at the Greville points. The order of these coefficients will be given by the monotonicity of the function  $f$ . Then we use the previous lemma to conclude. ■

### Preserving the convexity

**Lemma 4** If  $g = \sum_{j=0}^n c_j N_j^p$  is an element of the Schoenberg space  $\mathcal{S}_p(T)$ , then  $g$  is convex if

$$\nabla c_j \leq \nabla c_{j+1} \quad \forall 0 \leq j \leq n-1$$

*Proof.* We shall consider two cases

1.  $g$  is differentiable everywhere.

Then we know that

$$g'(x) = \sum_j \nabla c_j N_j^{p-1}(x)$$

But since  $\nabla c_j \leq \nabla c_{j+1} \quad \forall 0 \leq j \leq n-1$ , and using the lemma 3, we get that  $g'$  is an increasing function, which means that  $g$  is convex.

2. Now assume there exists only one point  $z$  where  $g$  is not differentiable and let  $x_0 < x_1 < x_2$  be three points in  $[a, b]$ . We shall prove that

$$\frac{g(x_1) - g(x_0)}{x_1 - x_0} \leq \frac{g(x_2) - g(x_1)}{x_2 - x_1}$$

The case where  $z \notin (x_0, x_2)$  is covered by the previous point. The interesting cases are the one for which  $z \in [x_0, x_2]$ . Assume for the moment that  $x_0 < z < x_1$ . Then using the fact that

$$\frac{g(x_1) - g(x_0)}{x_1 - x_0} = \frac{g(z) - g(x_0)}{z - x_0} \frac{z - x_0}{x_1 - x_0} + \frac{g(x_1) - g(z)}{x_1 - z} \frac{x_1 - z}{x_1 - x_0}$$

we get

$$\frac{g(x_1) - g(x_0)}{x_1 - x_0} \leq \frac{g(x_1) - g(z)}{x_1 - z}$$

but since  $g$  is convex on the right of  $z$ , we have also

$$\frac{g(x_1) - g(z)}{x_1 - z} \leq \frac{g(x_2) - g(x_1)}{x_2 - x_1}$$

then

$$\frac{g(x_1) - g(x_0)}{x_1 - x_0} \leq \frac{g(x_1) - g(z)}{x_1 - z} \leq \frac{g(x_2) - g(x_1)}{x_2 - x_1}$$

Now if  $z = x_1$ , we use the mean value theorem and the fact that  $f'$  is increasing at the interior of each sub-interval.

The other cases and the one with several discontinuities is treated in a similar way. ■

**Proposition 9.2.4** If  $f \in \mathcal{C}[a, b]$  is convex then  $Vf$  is convex.

*Proof.* If  $f$  is convex then we have

$$\frac{f(t_j^*) - f(t_{j-1}^*)}{t_j^* - t_{j-1}^*} \leq \frac{f(t_{j+1}^*) - f(t_j^*)}{t_{j+1}^* - t_j^*}$$

$$\nabla c_j \leq \nabla c_{j+1} \quad \forall 0 \leq j \leq n-1$$

where  $c_j = f(t_j^*)$ . We conclude by using the previous lemma. ■

### 9.3 Quasi-Interpolation

In this section, we give a general procedure to construct quasi-interpolant for B-Splines and we provide some examples.

#### 9.3.1 General recipe for quasi-interpolants

In the sequel, we shall give a general procedure to construct local approximation on a Schoenberg space. Let  $f \in \mathcal{C}[a, b]$ . We will denote the quasi-interpolant by  $\mathcal{Q}$  and assume that

$$\mathcal{Q}f = \sum_{j=0}^n \lambda_j(f) N_j^p \quad (9.6)$$

1. We consider a span index  $j$ ,
2. We assume that there exist  $\mu$  and  $\nu$  such that the interval  $I := [t_\mu, t_\nu] \subset [a, b]$  and the interior of  $I \cap [t_j, t_{j+p+1}]$  is not empty.
3. We choose a local approximation method  $P^l$  and determine an approximation  $P^l f$  to  $f$  on  $I$  such that

$$P^l f = \sum_{i=\mu-p}^{\nu-1} b_i N_i^p \quad (9.7)$$

4. We set the coefficient  $\lambda_j f$  in Eq. 9.6 to  $b_j$

The last point makes sense only if  $\mu - p \leq j \leq \nu - 1$ , which is true since it is equivalent to having  $j + 1 \leq \nu$  and  $\nu \leq j + p$ . But this is simply a consequence of the second point.

It is important to notice that the global representation of polynomials and splines are just a consequence of the local case. The following lemma states this idea.

**Lemma 5** If  $\mathcal{Q}$  is constructed using the previous procedure, then for an integer  $l \leq p$ , if  $P^l$  reproduces the polynomial space  $\Pi_l(I)$  the quasi-interpolant  $\mathcal{Q}$  reproduces also all polynomials of degree  $l$  on  $[a, b]$ , *i.e.*

$$P^l g = g \quad \forall g \in \Pi_l(I) \quad \Rightarrow \quad \mathcal{Q}g = g \quad \forall g \in \Pi_l \quad (9.8)$$

Moreover, if  $P^l$  reproduces the splines on  $I$  then it reproduces the splines on  $[a, b]$ , *i.e.*

$$P^l g = g \quad \forall g \in \mathcal{S}_p(T, I) \quad \Rightarrow \quad \mathcal{Q}g = g \quad \forall g \in \mathcal{S}_p(T) \quad (9.9)$$

*Proof.* TODO ■

#### 9.3.2 Examples

In the following examples, we assume that all interior knots are distincts.

##### Piecewise linear interpolation

In this case, we can define  $I := [t_j, t_{j+1}]$  and

$$P^l f = f(t_j) N_{j-1}^1 + f(t_{j+1}) N_j^1$$

which gives the global approximation

$$\mathcal{Q}f = \sum_j f(t_j) N_j^1 \quad (9.10)$$

### 3-point quadratic quasi-interpolation

We consider  $I := [t_{j+1}, t_{j+2}]$  and the local spline space  $\mathcal{S}_2(T, I) := \text{span}(N_i^2, j-1 \leq i \leq j+1)$ . On  $I$  we take three points

$$\begin{cases} x_0^j = t_{j+1} \\ x_1^j = \frac{t_{j+1} + t_{j+2}}{2} \\ x_2^j = t_{j+2} \end{cases}$$

The local approximation is defined as

$$P^I f = \sum_{i=j-1}^{j+1} b_i N_i^2 \quad (9.11)$$

In order to find the coefficients  $b_i$ , we need to solve the linear system

$$P^I f(x_k^j) = f(x_k^j), \quad k \in \{0, 1, 2\} \quad (9.12)$$

In the next section, we shall show how to compute these coefficients. For the moment, we will give the global approximation in terms of the functionals  $\lambda_j$

$$\lambda_j f = \begin{cases} f(t_0) & , j = 0 \\ -\frac{1}{2}f(x_0^j) + 2f(x_1^j) - \frac{1}{2}f(x_2^j) & , 1 \leq j \leq n-1 \\ f(t_{n+1}) & , j = n \end{cases} \quad (9.13)$$

### 5-point cubic quasi-interpolation

We consider  $I := [t_{j+1}, t_{j+3}]$  and the local spline space  $\mathcal{S}_3(T, I) := \text{span}(N_i^3, j-2 \leq i \leq j+2)$ . On  $I$  we take five points

$$\begin{cases} x_0^j = t_{j+1} \\ x_1^j = \frac{t_{j+1} + t_{j+2}}{2} \\ x_2^j = t_{j+2} \\ x_3^j = \frac{t_{j+2} + t_{j+3}}{2} \\ x_4^j = t_{j+3} \end{cases}$$

The local approximation is defined as

$$P^I f = \sum_{i=j-2}^{j+2} b_i N_i^3 \quad (9.14)$$

In order to find the coefficients  $b_i$ , we need to solve the linear system

$$P^I f(x_k^j) = f(x_k^j), \quad k \in \{0, 1, 2, 3, 4\} \quad (9.15)$$

The global approximation in terms of the functionals  $\lambda_j$

$$\lambda_j f = \begin{cases} f(t_3) & , j = 0 \\ -\frac{1}{18} \left( -5f(x_0^j) + 40f(x_1^j) - 24f(x_2^j) + 8f(x_3^j) - f(x_4^j) \right) & , j = 1 \\ -\frac{1}{6} \left( f(x_0^j) - 8f(x_1^j) + 20f(x_2^j) - 8f(x_3^j) + f(x_4^j) \right) & , 2 \leq j \leq n-2 \\ -\frac{1}{18} \left( -f(x_0^j) + 8f(x_1^j) - 24f(x_2^j) + 40f(x_3^j) - 5f(x_4^j) \right) & , j = n-1 \\ f(t_{n+1}) & , j = n \end{cases} \quad (9.16)$$

### 9.3.3 Computing the coefficients functionals

In the sequel, we shall explain how to compute the coefficients  $\lambda_j$  through the example of the 3-point quadratic quasi-interpolant.

We will be looking for  $\lambda_j$  in the general form

$$\lambda_j : f \mapsto \omega_0 f(x_0^j) + \omega_1 f(x_1^j) + \omega_2 f(x_2^j)$$

By taking  $f$  to be one of the B-Splines  $\{N_{j-1}^2, N_j^2, N_{j+1}^2\}$  we get the linear system

$$\begin{cases} \lambda_j N_{j-1}^2 &= \omega_0 N_{j-1}^2(x_0^j) + \omega_1 N_{j-1}^2(x_1^j) + \omega_2 N_{j-1}^2(x_2^j) \\ \lambda_j N_j^2 &= \omega_0 N_j^2(x_0^j) + \omega_1 N_j^2(x_1^j) + \omega_2 N_j^2(x_2^j) \\ \lambda_j N_{j+1}^2 &= \omega_0 N_{j+1}^2(x_0^j) + \omega_1 N_{j+1}^2(x_1^j) + \omega_2 N_{j+1}^2(x_2^j) \end{cases} \quad (9.17)$$

On the other hand, by construction we have  $\lambda_j N_i^2 = \delta_{ij}$ . Therefore, the left hand side of our system is known. Now what remains is to have numerical values for the B-Splines evaluation on  $\{x_0^j, x_1^j, x_2^j\}$ . Since  $I = [t_{j+1}, t_{j+2}]$  only the knots  $t_{j+1}$  and  $t_{j+2}$  are *seen* by the space of quadratic polynomials on  $I$ . We can then extend these knots with any knots we want to get a set of B-Splines that can be easily evaluated (let's say analytically). The simplest way is to consider the Bernstein polynomials, meaning we take the knot sequence  $\{t_{j+1}, t_{j+1}, t_{j+1}, t_{j+2}, t_{j+2}, t_{j+2}\}$ . In this case, we get the linear system

$$\begin{cases} \omega_0 B_0^2(0) + \omega_1 B_0^2(\frac{1}{2}) + \omega_2 B_0^2(1) = 0 \\ \omega_0 B_1^2(0) + \omega_1 B_1^2(\frac{1}{2}) + \omega_2 B_1^2(1) = 1 \\ \omega_0 B_2^2(0) + \omega_1 B_2^2(\frac{1}{2}) + \omega_2 B_2^2(1) = 0 \end{cases} \quad (9.18)$$

which gives

$$\begin{cases} 0 &= \omega_0 + \frac{1}{4}\omega_1 \\ 1 &= \frac{1}{2}\omega_1 \\ 0 &= \frac{1}{4}\omega_1 + \omega_2 \end{cases} \quad (9.19)$$

Therefore

$$\begin{cases} \omega_0 &= -\frac{1}{2} \\ \omega_1 &= 2 \\ \omega_2 &= -\frac{1}{2} \end{cases} \quad (9.20)$$

## 9.4 Global approximation

### 9.4.1 Interpolation

Given the set of interpolation points  $X := \{x_0, \dots, x_n\}$  and data  $Y := \{y_0, \dots, y_n\}$ , we aim to find a spline  $s$  such that  $y_i = s(x_i)$ . The Spline interpolation problem is

**Definition 9.4.1 — Spline interpolation.** Find a spline  $s := \sum_{j=0}^n c_j N_j^p \in \mathcal{S}_p(T)$  such that

$$s(x_i) = y_i \quad 0 \leq i \leq n \quad (9.21)$$

In a matrix form, the Spline interpolation problem writes

$$Mc = y \quad (9.22)$$

where  $c$  is the unknown vector of the spline coefficients,

$$c = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$M$  is the **collocation matrix** given by

$$M = \begin{pmatrix} N_0^p(x_0) & \dots & N_n^p(x_0) \\ N_0^p(x_1) & \dots & N_n^p(x_1) \\ \vdots & \dots & \vdots \\ N_0^p(x_n) & \dots & N_n^p(x_n) \end{pmatrix} \quad (9.23)$$

while  $y$  is the given data

$$y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Notice that when interpolating a function  $f$ , the given data will be  $y_i := f(x_i)$ .

In general the linear system given by the Eq. 9.22 is not always solvable. The following result by Whitney-Schoenberg gives a necessary and sufficient condition to ensure that the interpolation problem has a unique solution.

**Theorem 9.4.1** The collocation matrix is nonsingular if and only if the diagonal elements are positive, *i.e.*

$$N_i^p(x_i) > 0 \quad \forall i \in \{0, \dots, n\} \quad (9.24)$$

This condition is also equivalent to

$$t_i < x_i < t_{i+p+1} \quad \forall i \in \{0, \dots, n\} \quad (9.25)$$

**Exercise 9.2** Show that the Greville points fulfill the Whitney-Schoenberg condition Eq. 9.24.

### 9.4.2 Histopolation

Another interesting way to approximate a function, is to preserve the integrals between given points, rather than the value of the function on these points. Given the set of interpolations points  $X := \{x_0, \dots, x_{n+1}\}$  and a continuous function  $f$ , the histopolation problem writes

**Definition 9.4.2 — Spline histopolation.** Find a spline  $s := \sum_{j=0}^n c_j N_j^p \in \mathcal{S}_p(T)$  such that

$$\int_{x_i}^{x_{i+1}} s \, dx = \int_{x_i}^{x_{i+1}} f \, dx \quad 0 \leq i \leq n \quad (9.26)$$

In a matrix form, the Spline histopolation problem writes

$$Mc = y \quad (9.27)$$

where  $c$  is the unknown vector of the spline coefficients,

$$c = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$M$  is the **histopolation matrix** given by

$$M = \begin{pmatrix} \int_{x_0}^{x_1} N_0^p \, dx & \dots & \int_{x_0}^{x_1} N_n^p \, dx \\ \int_{x_1}^{x_2} N_0^p \, dx & \dots & \int_{x_1}^{x_2} N_n^p \, dx \\ \vdots & \dots & \vdots \\ \int_{x_n}^{x_{n+1}} N_0^p \, dx & \dots & \int_{x_n}^{x_{n+1}} N_n^p \, dx \end{pmatrix} \quad (9.28)$$

while  $y$  is the given data

$$y = \begin{pmatrix} \int_{x_0}^{x_1} f \, dx \\ \int_{x_1}^{x_2} f \, dx \\ \vdots \\ \int_{x_n}^{x_{n+1}} f \, dx \end{pmatrix}$$

**Theorem 9.4.2** The histopolation matrix is nonsingular if and only if

$$t_i < x_i < t_{i+p+1} \quad \forall i \in \{0, \dots, n\} \quad (9.29)$$

*Proof.* TODO ■

### Histopolation using M-Splines

Rather than using the B-Splines for the histopolation problem, one can use the M-Splines. In this case, the histopolation matrix given in Eq. 9.28 can be computed easily using the following result.

**Proposition 9.4.3** For every  $0 \leq i \leq n$  and  $0 \leq j \leq n$ , we have

$$\int_{x_i}^{x_{i+1}} M_j^p(t) dt = \sum_{k=0}^{j-1} (N_k^p(x_i) - N_k^p(x_{i+1})) \quad (9.30)$$

*Proof.* Integrating the relation  $\frac{d}{dt}N_k^p(t) = M_k^p(t) - M_{k+1}^p(t)$  on the interval  $[x_i, x_{i+1}]$ , we have

$$N_k^p(x_{i+1}) - N_k^p(x_i) = \int_{x_i}^{x_{i+1}} (M_k^p(t) - M_{k+1}^p(t)) dt$$

summing the last equation for  $k = 0$  to  $k = j - 1$ , we get

$$\begin{aligned} \sum_{k=0}^{j-1} (N_k^p(x_{i+1}) - N_k^p(x_i)) &= \int_{x_i}^{x_{i+1}} \sum_{k=0}^{j-1} (M_k^p(t) - M_{k+1}^p(t)) dt \\ &= \int_{x_i}^{x_{i+1}} (M_0^p(t) - M_j^p(t)) dt \end{aligned}$$

hence,

$$\int_{x_i}^{x_{i+1}} M_j^p(t) dt = \sum_{k=0}^{j-1} (N_k^p(x_i) - N_k^p(x_{i+1}))$$

■

**R** The last result gives an optimized implementation for the assembly of the histogram matrix, since the right hand side term can be computed by accumulating the summation for each  $j$ .

### 9.4.3 Least-square approximation

In the sequel, we consider a set of points  $X := \{x_0, \dots, x_m\}$  where  $m \geq n$  and given data  $Y := \{y_0, \dots, y_m\}$ . The least square problem writes

**Definition 9.4.3 — Spline least-square approximation.** Find a spline  $s := \sum_{j=0}^n c_j N_j^p \in \mathcal{S}_p(T)$  such that

$$\min_{g \in \mathcal{S}_p(T)} \sum_{i=0}^m (g(x_i) - y_i)^2 \quad (9.31)$$

where  $(w_i)_{i=0}^m$  is a set of positive weights.

**Lemma 6** The problem 9.31 is equivalent to the linear least square problem

$$\min_{c \in \mathbb{R}^n} \|Mc - b\|^2 \quad (9.32)$$

where  $M_{ij} = \sqrt{w_i} N_j^p(x_i)$  and  $b_i = \sqrt{w_i} y_i$ .

*Proof.* TODO

■

**Lemma 7** The matrix  $M^T M$  is symmetric and nonsingular.

*Proof.* TODO

■



**Proposition 9.4.4** The problem 9.32 has a solution given by

$$M^T M c = M^T b \quad (9.33)$$

The solution is unique if  $M$  has linearly independent columns.

*Proof.* TODO ■

**Theorem 9.4.5** The problem 9.32 has a unique solution if and only if we can find a sub-sequence  $(x_{i_l})_{l=0}^n$  such that

$$N_l^p(x_{i_l}) \neq 0 \quad \forall l \in \{0, \dots, n\} \quad (9.34)$$

*Proof.* TODO ■

## 9.5 Approximation with Quasi-Interpolation

In this section, we are interested in local linear methods to construct an approximation of a given function  $f$ . This means that each functional  $\lambda_i$  will be a linear functional that depends only on the values of  $f$  on the support of  $N_i^p$ .

In the sequel, we will assume that  $f \in \mathcal{C}^{-1}[a, b]$  and we consider a knot vector such that  $t_p = a$  and  $t_{n+1} = b$ . We first start by giving some definitions.

**Definition 9.5.1** Let  $S \subset [a, b]$  be a nonempty set. A linear functional  $\lambda : \mathcal{C}^{-1}[a, b] \mapsto \mathbb{R}$  is said to be **supported** on  $S$  if

$$\lambda(f) = 0, \quad \forall f \in \mathcal{C}^{-1}[a, b] \quad \text{such that } f|_S = 0 \quad (9.35)$$

**Definition 9.5.2** The quasi-interpolant  $\mathcal{Q}$  is called a **local quasi-interpolant** if

1. each  $\lambda_i$  is supported on the interval  $I_i$ , where

$$I_i := [t_i, t_{i+p+1}] \cap [a, b] \quad (9.36)$$

2. the quasi-interpolant  $\mathcal{Q}$  reproduces  $\Pi_l$ , for some  $l \in [0, p]$ , i.e.

$$\mathcal{Q}f(x) = f(x), \quad \forall x \in [a, b] \text{ and } \forall f \in \Pi_l \quad (9.37)$$

**Definition 9.5.3** A local quasi-interpolant  $\mathcal{Q}$  is **bounded** in a  $L_q$ -norm, with  $1 \leq q \leq \infty$ , if there is a constant  $C_{\mathcal{Q}}$  such that for each  $\lambda_i$  we have

$$\forall f \in \mathcal{C}^{-1}[a, b], \quad |\lambda_i f| \leq C_{\mathcal{Q}} h_i^{-\frac{1}{q}} \|f\|_{L_q(I_i)} \quad (9.38)$$

with  $h_i = \max_{l \in D_i} t_{l+1} - t_l$ , with  $D_i := [\max(i, p+1), \min(i+p, n)]$ .

**Lemma 8** Suppose that  $f \in L_q([t_p, t_{n+1}])$  for some  $q$ , with  $1 \leq q \leq \infty$  and there exists a set of strictly increasing integers  $\{m_{i_1}, \dots, m_{i_2}\}$ , with  $t_p \leq t_{m_{i_1}}$  and  $t_{m_{i_2}+r} \leq t_{n+1}$  for some positive integer  $r$  and integers  $i_1$  and  $i_2$ . Then

$$\left( \sum_{j=i_1}^{i_2} \|f\|_{L_q([t_{m_j}, t_{m_j+r}])}^q \right)^{\frac{1}{q}} \leq r^{\frac{1}{q}} \|f\|_{L_q([t_p, t_{n+1}])} \quad (9.39)$$

*Proof.* TODO ■

**Theorem 9.5.1 — Local estimation.** Let  $\mathcal{Q}$  be a bounded local quasi-interpolant in an  $L_q$ -norm, with  $1 \leq q \leq \infty$ . Let  $l$  and  $p$  be integers with  $0 \leq l \leq p$ . Suppose  $t_m < t_{m+1}$  for some  $p \leq m \leq n$  and let  $f \in W_q^{l+1}(J_m)$  with

$$J_m := [t_{m-p}, t_{m+p+1}] \cap [a, b]$$

Then

$$\|f - \mathcal{Q}f\|_{L_q([t_m, t_{m+1}])} \leq \frac{(2p+1)^{l+1}}{l!} (1 + C_{\mathcal{Q}}) h_m^{l+1} \|f^{(l+1)}\|_{L_q(J_m)} \quad (9.40)$$

where  $h_m = \max_{m-p \leq i \leq m+p, [t_i, t_{i+1}] \subset J_m} t_{i+1} - t_i$

*Proof.* We recall that for  $l \geq 0$ ,  $f$  is a continuous function. Let  $x \in [t_m, t_{m+1})$ . Using lemma 2 and the fact that  $\mathcal{Q}$  is bounded in  $L_q$ -norm, we have

$$|\mathcal{Q}f(x)| \leq \max_{m-p \leq j \leq m} |\lambda_j f| \leq \max_{m-p \leq j \leq m} C_{\mathcal{Q}} h_j^{-\frac{1}{q}} \|f\|_{L_q(I_j)}$$

But  $t_{m+1} - t_m \leq \max_{m-p \leq j \leq m} h_j$  and  $J_m = \bigcup_{m-p \leq j \leq m} I_j$ , we get

$$\|\mathcal{Q}f\|_{L_q([t_m, t_{m+1}])} \leq C_{\mathcal{Q}} \|f\|_{L_q(J_m)} \quad (9.41)$$

Now let  $g$  be any polynomial of degree  $l$ , i.e.  $g \in \Pi_l$ . Since  $\mathcal{Q}$  reproduces all polynomials of degree  $l$ , we have

$$\|f - \mathcal{Q}f\|_{L_q([t_m, t_{m+1}])} \leq \|f - g\|_{L_q([t_m, t_{m+1}])} + \|\mathcal{Q}(g - f)\|_{L_q([t_m, t_{m+1}])}$$

But since  $[t_m, t_{m+1}] \subset J_m$ , using Eq. 9.41 we get

$$\|f - \mathcal{Q}f\|_{L_q([t_m, t_{m+1}])} (1 + C_{\mathcal{Q}}) \|f - g\|_{L_q(J_m)}$$

Now let us choose  $g$  to be the Taylor polynomial of degree  $l$  with  $\tau = \max(t_{m-p}, a)$ . By taking  $r = 0$  in refTODO we get

$$\|f - g\|_{L_q(J_m)} \leq \frac{(2p+1)^{l+1}}{l!} h_m^{l+1} \|f^{(l+1)}\|_{L_q(J_m)} \quad (9.42)$$

Therefore

$$\|f - \mathcal{Q}f\|_{L_q([t_m, t_{m+1}])} (1 + C_{\mathcal{Q}}) \frac{(2p+1)^{l+1}}{l!} h_m^{l+1} \|f^{(l+1)}\|_{L_q(J_m)}$$

■

**Theorem 9.5.2 — Global estimation.** Let  $\mathcal{Q}$  be a bounded local quasi-interpolant in an  $L_q$ -norm, with  $1 \leq q \leq \infty$ . Let  $l$  and  $p$  be integers with  $0 \leq l \leq p$ . For all  $f \in W_q^{l+1}([a, b])$  we have

$$\|f - \mathcal{Q}f\|_{L_q([a, b])} \leq \frac{(2p+1)^{l+1+\frac{1}{q}}}{l!} (1 + C_{\mathcal{Q}}) h^{l+1} \|f^{(l+1)}\|_{L_q([a, b])} \quad (9.43)$$

where  $h = \max_{p \leq i \leq n} t_{i+1} - t_i$

*Proof.* Use the local estimation theorem and the lemma 8 gives the desired result. ■

### 9.5.1 Quasi-interpolation and reproduction of polynomials

In order to use the previous results, one still needs to show that a quasi-interpolant reproduces polynomials of a given degree  $l$ . In the sequel, we shall give conditions to find the degree  $l$ .

**Proposition 9.5.3** We consider a set of  $n+1$  basis functions for  $\Pi_l$ , given by

$$\{\phi_{j,0}, \dots, \phi_{j,l}\}, \quad j \in \{0, \dots, n\}, \quad 0 \leq l \leq p \quad (9.44)$$

and we consider their B-Splines representations as

$$\phi_{i,j} = \sum_{k=0}^n c_{ijk} N_k^p \quad (9.45)$$

A linear quasi-interpolant  $\mathcal{Q}$  reproduces  $\Pi_l$  if the corresponding linear functionals satisfy

$$\lambda_j(\phi_{i,j}) = c_{jij}, \quad j \in \{0, \dots, n\}, \quad i \in \{0, \dots, l\} \quad (9.46)$$

*Proof.* TODO ■

The next proposition gives a sufficient condition for a quasi-interpolant to reproduce the Schoenberg space.

**Proposition 9.5.4** The linear quasi-interpolant  $\mathcal{Q}$  reproduces the whole spline space if

1.  $\mathcal{Q}$  reproduces  $\Pi_l$
2. each linear functional  $\lambda_j$  is supported on one knot interval

$$[t_{m_j}^+, t_{m_{j+1}}^-] \subset [t_j, t_{j+p+1}] \quad (9.47)$$

*Proof.* TODO ■

## 9.5.2 Examples

### Variation diminishing approximation

**Proposition 9.5.5** For every function  $f \in W_\infty^2([a, b])$  we have

$$\|f - Vf\|_{L_\infty([a, b])} \leq 2(2p+1)^2 h^2 \|f''\|_{L_\infty([a, b])} \quad (9.48)$$

**Exercise 9.3** Prove the last proposition.

(Hint: use the global estimation theorem with the  $L_\infty$ -norm.) ■

### 3-points quadratic quasi-interpolation

**Proposition 9.5.6** For every function  $f \in W_\infty^3([a, b])$  we have

$$\|f - \mathcal{Q}_2 f\|_{L_\infty([a, b])} \leq 250h^3 \|f^{(3)}\|_{L_\infty([a, b])} \quad (9.49)$$

*Proof.* TODO ■

## 9.6 Approximation power of Splines

The following result gives the approximation power on the Schoenberg space. After stating this important result, we shall give the proof through different steps.

**Proposition 9.6.1 — Approximation Error.** Let  $f \in W_q^{l+1}([a, b])$  with  $1 \leq q \leq \infty$  and  $0 \leq l \leq p$ . Then there exists a spline  $s_p \in \mathcal{S}_p(T)$  such that

$$\|D^{(r)}(f - s_p)\|_{L_q([a, b])} \leq Kh^{l+1-r} \|f^{(l+1)}\|_{L_q([a, b])}, \quad 0 \leq r \leq l \quad (9.50)$$

where  $h := \max_{p \leq j \leq n}$  and  $K$  is a constant depending only on  $p$ .

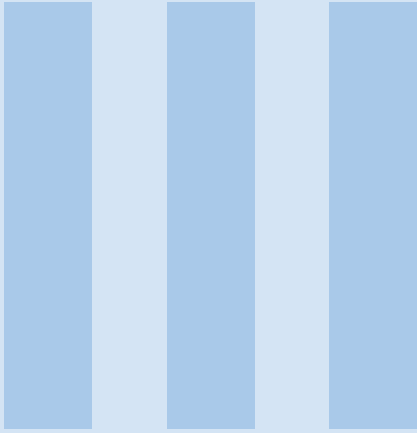
**R** The constant  $K$  grows exponentially with  $p$ , but can be removed in some cases.

## 9.7 Problems

TODO

## 10. Historical Notes

TODO



# Finite Elements method

<b>11</b>	<b>Functional Analysis</b> .....	<b>96</b>
11.1	Notations and Preliminaries	
11.2	Sobolev spaces	
11.3	The Sobolev space $H(\text{curl}, \Omega)$	
11.4	The Sobolev space $H(\text{div}, \Omega)$	
11.5	DeRham sequences	
11.6	Problems	
<b>12</b>	<b>Galerkin methods</b> .....	<b>101</b>
12.1	Abstract framework	
12.2	Galerkin Approximation	
12.3	Saddle-point problems	
12.4	Problems	
<b>13</b>	<b>Historical Notes</b> .....	<b>115</b>



# 11. Functional Analysis

## 11.1 Notations and Preliminaries

In the course of the lecture we shall work with the Sobolev spaces  $H^m(\Omega)$ ,  $H(\operatorname{div}, \Omega)$  and  $H(\operatorname{curl}, \Omega)$  and recall here their basic properties without proof. For a more detailed presentation with proofs we refer to Section 2.1. of [1].

- we use boldface notation for spaces of vector functions. For instance in 3D,  $\mathbf{H}^1(\Omega)$  denotes the space  $(H^1(\Omega))^3$ .
- for an operator  $d \in \{\nabla, \nabla \times, \nabla \cdot\}$ , we denote the kernel space  $\mathcal{N}(d)$  while the range will be  $\mathcal{R}(d)$ .



We shall need the following Sobolev spaces, given first without boundary conditions,

$$H^1(\Omega) = \{\varphi \in L^2(\Omega), \nabla\varphi \in L^2(\Omega)\} \quad (11.1)$$

$$\mathbf{H}^1(\Omega) = \{\Psi \in L^2(\Omega), \nabla\Psi \in L^2(\Omega)\} \quad (11.2)$$

$$\mathbf{H}(\text{curl}, \Omega) = \{\Psi \in L^2(\Omega), \nabla \times \Psi \in L^2(\Omega)\} \quad (11.3)$$

$$\mathbf{H}(\text{div}, \Omega) = \{\Psi \in L^2(\Omega), \nabla \cdot \Psi \in L^2(\Omega)\} \quad (11.4)$$

and using the correspondant boundary conditions,

$$H_0^1(\Omega) = \{\varphi \in H^1(\Omega), \varphi = 0 \text{ on } \partial\Omega\} \quad (11.5)$$

$$\mathbf{H}_0^1(\Omega) = \{\Psi \in \mathbf{H}^1(\Omega), \Psi = 0 \text{ on } \partial\Omega\} \quad (11.6)$$

$$\mathbf{H}_0(\text{curl}, \Omega) = \{\Psi \in \mathbf{H}(\text{curl}, \Omega), \Psi \times \mathbf{n} = 0 \text{ on } \partial\Omega\} \quad (11.7)$$

$$\mathbf{H}_0(\text{div}, \Omega) = \{\Psi \in \mathbf{H}(\text{div}, \Omega), \Psi \cdot \mathbf{n} = 0 \text{ on } \partial\Omega\} \quad (11.8)$$

$$L_0^2(\Omega) = \left\{ \varphi \in L^2(\Omega); \int_{\Omega} \varphi = 0 \right\} \quad (11.9)$$

Scalar-valued test functions will be denoted by  $\varphi$ , while  $\varphi_i$  will denote a scalar basis function, after reordering all the basis functions of a given discrete space.

Vector-valued functions will be written in bold, like  $\mathbf{u}, \mathbf{v}$ .  $\Psi$  will denote a vector-valued test function, while  $\Psi_i$  will denote a vector-valued basis function (after reordering the basis functions).

Even if most of what follows holds for both the  $2D$  and  $3D$  cases, we will restrict our studies to the  $2D$  one. We recall that in  $2D$ , there are two *curl* operators, one acting on scalars  $\nabla \times \phi = (\partial_y \phi, -\partial_x \phi)$  and one acting on vectors  $\nabla \times \Psi = \partial_x \Psi^y - \partial_y \Psi^x$ . Differential operators that return a vector (grad, curl) will be written in bold ( $\nabla, \nabla \times$ ). We also recall the Green formula for the *divergence* and *curl/rotational* operators

$$\int_{\Omega} (\nabla \cdot \mathbf{F}) G = - \int_{\Omega} \mathbf{F} \cdot \nabla G + \int_{\partial\Omega} (\mathbf{F} \cdot \mathbf{n}) G, \quad \forall \mathbf{F} \in \mathbf{H}(\text{div}, \Omega), \forall G \in H^1(\Omega) \quad (11.10)$$

$$\int_{\Omega} (\nabla \times G) \cdot \mathbf{F} = \int_{\Omega} G \nabla \times \mathbf{F} - \int_{\partial\Omega} (G \times \mathbf{n}) \cdot \mathbf{F}, \quad \forall \mathbf{F} \in \mathbf{H}(\text{curl}, \Omega), \forall G \in H^1(\Omega) \quad (11.11)$$

If  $\Omega \subset \mathbb{R}^d$ ,  $\mathbf{u} = (u_1, u_2, \dots, u_d)$  and  $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_d)$ , we recall the notation

$$\int_{\Omega} \nabla \mathbf{u} : \nabla \Psi = \sum_{i=1}^d \int_{\Omega} \nabla u_i \cdot \nabla \Psi_i \quad (11.12)$$

## 11.2 Sobolev spaces

We shall denote by  $\mathcal{D}(\Omega)$  the space of distribution.

### 11.2.1 The Sobolev space $W^{s,m}$

We start by recalling the definition of Sobolev spaces:

**Definition 11.2.1** Let  $s$  and  $m$  be two integers with  $s \geq 0$  and  $1 \leq m \leq \infty$ . The Sobolev space  $W^{s,m}(\Omega)$  is defined as

$$W^{s,m}(\Omega) = \{u \in \mathcal{D}'(\Omega), D^\alpha u \in L^m(\Omega), |\alpha| \leq s\} \quad (11.13)$$

The space  $W^{s,m}(\Omega)$  can be equipped with norm

$$\|u\|_{W^{s,m}(\Omega)} := \sum_{|\alpha| \leq s} \|D^\alpha u\|_{L^p(\Omega)} \quad (11.14)$$

### 11.2.2 The Sobolev space $H^m$

For any integer  $m \geq 1$ , one can define

$$H^m(\Omega) = W^{m,2}(\Omega) := \{v \in L^2(\Omega) \mid D^\alpha v \in L^2(\Omega), |\alpha| \leq m\} \quad (11.15)$$

$H^m(\Omega)$  is a Hilbert space equipped with the scalar product

$$(u, v)_{m,\Omega} := \sum_{|\alpha| \leq m} \int_{\Omega} D^\alpha u D^\alpha v \quad (11.16)$$

the associated norm will be denoted  $\|\cdot\|_{s,\Omega}$ .

The most classical second order operator is the Laplace operator, which reads in an arbitrary dimension  $d$  (generally  $d = 1, 2$  or  $3$ ),

$$\Delta u = \sum_{i=1}^d \frac{\partial^2 u}{\partial x_i^2}.$$

The classical Green formula for the Laplace operator reads: for  $u \in H^1(\Omega)$  and  $v \in H^1(\Omega)$

$$-\int_{\Omega} \Delta u v \, dx = \int_{\Omega} \nabla u \cdot \nabla v \, dx - \int_{\partial\Omega} \frac{\partial u}{\partial n} v \, d\sigma. \quad (11.17)$$

For essential boundary conditions related with this Green formula we shall define the space

$$H_0^1(\Omega) = \{v \in H^1(\Omega) \mid v|_{\partial\Omega} = 0\}.$$

Another classical operator which comes from elasticity is the bilaplacian operator  $\Delta^2 = \Delta\Delta$ , which is a fourth order operator. The Green formula needed for variational formulations of PDEs based on the bilaplacian reads

$$\int_{\Omega} \Delta^2 u v \, dx = \int_{\Omega} u \Delta^2 v \, dx + \int_{\partial\Omega} \left( u \frac{\partial \Delta v}{\partial n} - v \frac{\partial \Delta u}{\partial n} + \Delta u \frac{\partial v}{\partial n} - \Delta v \frac{\partial u}{\partial n} \right) d\sigma \quad (11.18)$$

$$H_0^2(\Omega) = \{v \in H^1(\Omega) \mid v|_{\partial\Omega} = 0, \frac{\partial v}{\partial n}|_{\partial\Omega} = 0\}.$$

### 11.2.3 Inequalities

**Lemma 9 — Poincaré.** Let  $1 \leq p \leq \infty$  and  $\Omega$  be a bounded open set. Then, there exists a constant  $C = C(p, \Omega)$ , such that

$$\forall v \in W_0^{1,p}(\Omega), \quad C \|v\|_{L^p(\Omega)} \leq \|\nabla v\|_{L^p(\Omega)} \quad (11.19)$$

### 11.3 The Sobolev space $H(\text{curl}, \Omega)$

### 11.4 The Sobolev space $H(\text{div}, \Omega)$

### 11.5 DeRham sequences

For any function  $u \in H^1(\Omega)$  we have  $\nabla \times \nabla u = 0$ . On the other hand, we have for any function  $\mathbf{u} \in \mathbf{H}(\text{curl}, \Omega)$ ,  $\nabla \cdot \nabla \times \mathbf{u} = 0$ . We just have shown that  $\nabla(H^1(\Omega)) \subset \mathcal{N}(\nabla \times)$  and  $\nabla \times (\mathbf{H}(\text{curl}, \Omega)) \subset \mathcal{N}(\nabla \cdot)$ . This is summarized in the following diagram, known as DeRham sequence, without boundary conditions in this case,

$$\mathbb{R} \hookrightarrow H^1(\Omega) \xrightarrow{\nabla} \mathbf{H}(\text{curl}, \Omega) \xrightarrow{\nabla \times} \mathbf{H}(\text{div}, \Omega) \xrightarrow{\nabla \cdot} L^2(\Omega) \longrightarrow 0 \quad (11.20)$$

and using the correspondant boundary conditions,

$$H_0^1(\Omega) \xrightarrow{\nabla} \mathbf{H}_0(\text{curl}, \Omega) \xrightarrow{\nabla \times} \mathbf{H}_0(\text{div}, \Omega) \xrightarrow{\nabla \cdot} L_0^2(\Omega) \longrightarrow 0 \quad (11.21)$$

In fact, DeRham complexes are sequences of spaces  $V_i$  and operators  $d_i$  such that  $d_{i+1} \circ d_i = 0$ . It leads to a sepcific algebraic structure that has been subject to active research in Analysis and Algebraic Geometry. or in a differential forms setting

$$\mathbb{R} \hookrightarrow \Lambda^0 \xrightarrow{d} \Lambda^1 \xrightarrow{d} \Lambda^2 \xrightarrow{d} \Lambda^3 \longrightarrow 0 \quad (11.22)$$

where  $d$  stands for the exterior derivative, while  $\Lambda^k$  is the space of  $k$ -forms.

$$\begin{array}{ccccccc} H^1(\Omega) & \xrightarrow{\nabla} & \mathbf{H}(\text{curl}, \Omega) & \xrightarrow{\nabla \times} & \mathbf{H}(\text{div}, \Omega) & \xrightarrow{\nabla \cdot} & L^2(\Omega) \\ P_h^{\text{grad}} \downarrow & & P_h^{\text{curl}} \downarrow & & P_h^{\text{div}} \downarrow & & P_h^{L^2} \downarrow \\ V_h(\text{grad}, \Omega) & \xrightarrow{\nabla} & \mathbf{V}_h(\text{curl}, \Omega) & \xrightarrow{\nabla \times} & \mathbf{V}_h(\text{div}, \Omega) & \xrightarrow{\nabla \cdot} & V_h(L^2, \Omega) \end{array}$$

### 11.5.1 Exact discrete DeRham sequence

### 11.5.2 Space decompositions

More details can be found in [9, 15].

$$\mathcal{N}(\nabla \times) = \nabla H_0^1(\Omega) \oplus (\nabla H_0^1(\Omega))^\perp \quad (11.23)$$

where  $(\nabla H_0^1(\Omega))^\perp$  is the orthogonal of  $H_0^1(\Omega)$  in  $\mathbf{H}_0(\text{curl}, \Omega)$  with respect to its inner product. We denote this space  $K_N(\Omega)$ , which is refered as *the normal cohomology space*.

$$K_N(\Omega) := (\nabla H_0^1(\Omega))^\perp \quad (11.24)$$

$K_N(\Omega)$  is the space of harmonic functions which vanish on the exterior and are constant in the interior connected components of  $\partial\Omega$ . Note that  $K_N(\Omega) \subset \mathbf{H}_0^1(\Omega)$ .

**Theorem 11.5.1** If  $\mathbf{u} \in \mathbf{H}_0(\text{curl}, \Omega)$ , such that  $\nabla \times \mathbf{u} = 0$ , then there exists a unique scalar potential  $p \in H_0^1(\Omega)$  and a function  $\mathbf{f}_N \in K_N(\Omega)$  such that

$$\mathbf{u} = \nabla p + \mathbf{f}_N \quad (11.25)$$

**Theorem 11.5.2 — Helmholtz Decomposition.** For every  $\mathbf{u} \in L^2(\Omega)$  there exist a unique

- $p \in H_0^1(\Omega)$
- $\mathbf{f}_N \in K_N(\Omega)$
- $\mathbf{A} \in \{\mathbf{w} \in \mathbf{H}(\text{curl}, \Omega), \nabla \cdot \mathbf{w} = 0, \mathbf{n} \cdot \mathbf{w} = 0 \text{ } \partial\Omega, \langle \mathbf{w} \cdot \mathbf{n}, 1 \rangle_{\Gamma_i} = 0\}$

ensuring the following decomposition

$$\mathbf{u} = \nabla p + \nabla \times \mathbf{A} + \mathbf{f}_N \quad (11.26)$$

**R** If  $\Omega$  is homotopy equivalent to a ball, then

$$\mathcal{N}(\nabla \times) = \nabla H_0^1(\Omega) \quad (11.27)$$

Therefore, the Helmholtz decomposition writes

$$\mathbf{u} = \nabla p + \nabla \times \mathbf{A} \quad (11.28)$$

**Theorem 11.5.3 — Regular Decomposition of  $\mathbf{H}(\text{curl}, \Omega)$ .** For any  $\mathbf{u} \in \mathbf{H}(\text{curl}, \Omega)$  there exists a  $\mathbf{v} \in \mathbf{H}^1(\Omega)$  such that

1.  $\nabla \times \mathbf{v} = \nabla \times \mathbf{u}$
2.  $\|\mathbf{v}\|_{\mathbf{L}^2(\Omega)} \lesssim \|\mathbf{u}\|_{\mathbf{L}^2(\Omega)}$
3.  $\|\mathbf{v}\|_{\mathbf{H}^1(\Omega)} \lesssim \|\nabla \times \mathbf{u}\|_{\mathbf{L}^2(\Omega)}$

The following result can be found in [Pasciak2002a] and [Zhao2002a].

**Theorem 11.5.4 — Regular Decomposition of  $\mathbf{H}_0(\text{curl}, \Omega)$ .** For any  $\mathbf{u} \in \mathbf{H}_0(\text{curl}, \Omega)$  there exists a  $\mathbf{v} \in \mathbf{H}_0^1(\Omega)$  such that

1.  $\nabla \times \mathbf{v} = \nabla \times \mathbf{u}$
2.  $\|\mathbf{v}\|_{\mathbf{L}^2(\Omega)} \lesssim \|\mathbf{u}\|_{\mathbf{L}^2(\Omega)}$
3.  $\|\mathbf{v}\|_{\mathbf{H}_0^1(\Omega)} \lesssim \|\nabla \times \mathbf{u}\|_{\mathbf{L}^2(\Omega)}$

## 11.6 Problems

TODO

## 12. Galerkin methods

### 12.1 Abstract framework

We consider  $a$  and  $L$  to be continuous bilinear and linear forms, respectively, on a Hilbert space  $V$ . We want to find a computable approximation for the solution  $u \in V$  of the variational problem

$$a(u, v) = \langle L, v \rangle, \quad \forall v \in V \quad (12.1)$$

where  $\langle \cdot, \cdot \rangle$  denotes the duality product between  $V'$  and  $V$ .

The idea of Galerkin approximation, is to find the solution in a family of subspaces of finite dimension, then prove that the constructed solutions converge to the solution of the variational problem Eq. (12.1). There are two major strategies, the first one is based on the coercivity and the other one on the *inf-sup* conditions. While the coercivity is easy to use, unfortunately, most of problems in CFD do not fulfill it.

**Definition 12.1.1 — V-ellipticity or Coercivity.**  $a$  is said to be coercive, if there exists a constant  $\alpha > 0$  such that

$$a(v, v) \geq \alpha \|v\|_V^2, \quad \forall v \in V \quad (12.2)$$

**Definition 12.1.2 — *inf-sup* conditions.**  $a$  is said to satisfy the *inf-sup* conditions, if there exists a constant  $\alpha > 0$  such that

1.

$$\sup_{v \in V} \frac{a(u, v)}{\|v\|_V} \geq \alpha \|u\|_V, \quad \forall u \in V \quad (12.3)$$

2.

$$\sup_{u \in V} \frac{a(u, v)}{\|u\|_V} \geq \alpha \|v\|_V, \quad \forall v \in V \quad (12.4)$$

**R** Notice that when  $a$  satisfies the *inf-sup* conditions and it is symmetric, then both conditions are the same. In general, the two conditions can be written as

$$\inf_{u \in V} \sup_{v \in V} \frac{a(u, v)}{\|u\|_V \|v\|_V} > 0$$

and

$$\inf_{v \in V} \sup_{u \in V} \frac{a(u, v)}{\|u\|_V \|v\|_V} > 0$$

Finally, let us notice that the coercivity implies the *inf-sup* conditions.

**Lemma 10** If  $a$  is coercive then it satisfies the *inf-sup* conditions.

*Proof.* We have

$$\sup_{v \in V} \frac{a(u, v)}{\|v\|_V} \geq \frac{a(u, u)}{\|u\|_V}$$

then we conclude using the coercivity of  $a$ . ■

## 12.2 Galerkin Approximation

We consider a family of finite dimensional subspaces of  $V$ , denoted by  $(V_h)_{h>0}$ . The Galerkin approximation  $u_h \in V_h$  is defined as the solution of the variational problem Eq. (12.1) by restricting the test functions on  $V_h$ , *i.e.*

$$a(u_h, v) = \langle L, v \rangle, \quad \forall v \in V_h \tag{12.5}$$

It is important to notice that while coercivity is inherited on subspaces, the *inf-sup* conditions are not. It is therefore important to have an additional *inf-sup* condition on the subspace, known as **stability condition**.

### 12.2.1 Convergence under coercivity

We recall Cea's lemma, which states that the Galerkin approximation is bounded by the best approximation of  $u$  from the subspace.

**Lemma 11 — Cea.** If  $a$  is a continuous and coercive bilinear form, then

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v \in V_h} \|u - v\|_V \tag{12.6}$$

*Proof.* Since both  $u$  and  $u_h$  are solutions to the variational problem (12.1) respectively on  $V$  and  $V_h$ , we have

$$a(u - u_h, v) = 0, \quad \forall v \in V_h$$

therefor, for  $v \in V_h$ , we have

$$a(u - u_h, u - v) = a(u - u_h, u - u_h) + a(u - u_h, u_h - v) = a(u - u_h, u - u_h)$$

Using the coercivity, we have

$$\alpha \|u - u_h\|_V \leq a(u - u_h, u - u_h) = a(u - u_h, u - v)$$

finally, the continuity of  $a$  gives

$$a(u - u_h, u - v) \leq M \|u - u_h\|_V \|u - v\|_V$$

by combining the two previous inequalities we get the desired result. ■

**Theorem 12.2.1** If  $a$  is a continuous and coercive bilinear form and the subspaces  $V_h$  are such that

$$\lim_{h \rightarrow 0} d(u, V_h) = 0 \quad (12.7)$$

with  $d(u, V_h) := \inf_{v \in V_h} \|u - v\|_V$ . Then

$$\lim_{h \rightarrow 0} u_h = u$$

*Proof.* Follows immediatly from Cea's lemma. ■

### 12.2.2 Convergence under *inf-sup* conditions

As mentioned before, the *inf-sup* conditions are not inherited on the subspaces  $V_h$ . In general, we must prove that there exists  $\beta > 0$ , such that the *inf-sup* holds on  $V_h$ , *i.e.*

$$\sup_{v \in V_h} \frac{a(u, v)}{\|v\|_V} \geq \beta \|u\|_V, \quad \forall u \in V_h \quad (12.8)$$

Idealy,  $\beta$  should be independent of  $N$ , in order to get the convergence.

**Exercise 12.1** Prove that the second part of the *inf-sup* conditions is a consequence of the inequality (12.8). ■

As in the coercive case, we first state a result that compares the Galerkin approximation with the distance to the space of approximation. This result is a generalization of Cea's lemma, and is due to Babuska.

**Lemma 12 — Babuska.** If  $a$  is a continuous bilinear form and satisfies the *inf-sup+stability* conditions, then

$$\|u - u_h\|_V \leq \left(1 + \frac{M}{\beta}\right) \inf_{v \in V_h} \|u - v\|_V \quad (12.9)$$

*Proof.* Let  $v \in V_h$ . since  $v - u_h \in V_h$ , the stability condition gives

$$\sup_{\phi \in V_h} \frac{a(v - u_h, \phi)}{\|\phi\|_V} \geq \beta \|v - u_h\|_V$$

but  $a(u - u_h, v) = 0$ , meaning  $a(v - u_h, \phi) = a(v - u, \phi)$  for all  $\phi \in V_h$ . Now using the continuity of  $a$  we get

$$M \|v - u\|_V \geq \sup_{\phi \in V_h} \frac{a(v - u, \phi)}{\|\phi\|_V} \geq \beta \|v - u_h\|_V$$

we conclude the proof by using the triangle inequality

$$\|u - u_h\|_V \leq \|u - v\|_V + \|v - u_h\|_V \leq \|u - v\|_V + \frac{M}{\beta} \|v - u\|_V$$

■

**Theorem 12.2.2** If  $a$  is a continuous bilinear form and satisfies the *inf-sup+stability* conditions. If the subspaces  $V_h$  are such that condition (12.7) holds, then

$$\lim_{h \rightarrow 0} u_h = u$$

*Proof.* Follows immediatly from Babuska's lemma. ■

**R** Under the additional stability condition, we have  $\|u_h\|_V \leq \frac{1}{\beta} \|L\|_{V'}$ , which is valid uniformly in  $N$  if  $\beta$  is independent of  $N$ .

### 12.2.3 The three basic aspects of the Finite Elements method

Let  $\Omega \subset \mathbb{R}^d$ , with  $d \geq 1$ , be a bounded domain.

In order to apply the Galerkin method, we face, by definition the problem of constructing the family of finite dimensional subspaces  $V_h \subset V$ , such that  $V$  is  $H^1(\Omega)$ ,  $H_0^1(\Omega)$ ,  $H^2(\Omega)$ ,  $H(\text{curl}, \Omega)$ , ... As stating by P. Ciarlet, the Finite Elements Method is in its simplest form, a specific process of constructing the family  $(V_h)_{h \geq 0}$ . This construction is characterized by three basic aspects and are described below.

#### First basic aspect: Triangulation

A triangulation  $\mathcal{T}_h$  is established over  $\bar{\Omega}$ , i.e.  $\bar{\Omega}$  is subdivided into a finite number of subsets  $K$ , called **finite elements**, such that

1.  $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$
2. for all  $K \in \mathcal{T}_h$ ,  $K$  is closed and its interior is not empty
3. for all  $K_1 \neq K_2 \in \mathcal{T}_h$  we have  $\overset{\circ}{K}_1 \cap \overset{\circ}{K}_2 = \emptyset$
4. for all  $K \in \mathcal{T}_h$ ,  $\partial K$  is Lipschitz-continuous

#### Second basic aspect: power approximation

On every  $K \in \mathcal{T}_h$ , a space of functions  $P_K$  is constructed.  $P_K$  should contain polynomials or functions which are close to polynomials.

- this is the key to all convergence results
- it is also important for having simple and fast computations of the coefficients of the resulting linear system

#### Third basic aspect: basis functions

There exists at least one **canonical basis** in the space  $V_h$  whose corresponding basis functions have a *local support* property, are as small as possible and can be easily described. This aspect leads to sparsity in the resulting matrix.



### 12.2.4 Examples

#### Scalar linear elliptic equations of second order

For  $\Omega \subset \mathbb{R}^d$ , we consider the following problem

$$\begin{cases} - \sum_{1 \leq i, j \leq d} \partial_{x_i} (a_{ij} \partial_{x_j} u) = f, & \Omega \\ u = 0, & \partial\Omega \end{cases} \quad (12.10)$$

where the coefficients functions  $a_{ij} : \Omega \rightarrow \mathbb{R}$  are bounded and there exists  $\gamma > 0$  (ellipticity condition) such that

$$\gamma |y|^2 \leq \sum_{1 \leq i, j \leq d} a_{ij}(x) y_i y_j, \quad \forall x \in \Omega, \forall y \in \mathbb{R}^d \quad (12.11)$$

Before writing the variational formulation associated to (12.10), we need to define the space of function  $V$  which is in this case given by  $V := H_0^1(\Omega)$  where

$$H_0^1(\Omega) = \{u \in H^1(\Omega), u = 0 \text{ on } \partial\Omega\} \quad (12.12)$$

and

$$H^1(\Omega) = \{u \in L^2(\Omega), \nabla u \in (L^2(\Omega))^d\} \quad (12.13)$$

$H_0^1(\Omega)$  is a Hilbert space under the norm  $\|\cdot\|_{H_0^1(\Omega)}$  with

$$\|u\|_{H_0^1(\Omega)}^2 := \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 \quad (12.14)$$

the bilinear and linear forms are given by

$$a(u, v) := \sum_{1 \leq i, j \leq d} \int_{\Omega} a_{ij} \partial_{x_j} u \partial_{x_i} v \, dx$$

and

$$\langle L, v \rangle = \int_{\Omega} f v \, dx$$

Using the ellipticity condition, the boundness of the coefficients and the Poincaré inequality, we show that  $a$  is coercive and continuous. Moreover, if  $f \in L^2(\Omega)$  then  $L$  is continuous.

#### Biharmonic problem

We consider the following problem

$$\begin{cases} \Delta^2 u = f, & \Omega \\ u = 0, & \partial\Omega \\ \nabla u \cdot \mathbf{n} = 0, & \partial\Omega \end{cases} \quad (12.15)$$

which is known as the *homogeneous Dirichlet problem for the biharmonic operator*  $\Delta$ . The space of functions considered in this example  $V$  is given by  $V := H_0^2(\Omega)$  where

$$H_0^2(\Omega) = \{u \in H^2(\Omega), u = \nabla u \cdot \mathbf{n} = 0 \text{ on } \partial\Omega\} \quad (12.16)$$

and

$$H^2(\Omega) = \{u \in L^2(\Omega), \nabla u \in (L^2(\Omega))^d\} \quad (12.17)$$

$H_0^2(\Omega)$  is a Hilbert space under the norm  $v \mapsto |\Delta v|_{L^2(\Omega)}$  with the bilinear and linear forms are given by

$$a(u, v) := \int_{\Omega} \Delta u \Delta v \, dx$$

and

$$\langle L, v \rangle = \int_{\Omega} f v \, dx$$

**$H(\text{curl}, \Omega)$ -elliptic problem**

Let  $\Omega \subset \mathbb{R}^d$  be an open Lipschitz bounded set, and we look for the solution of the following problem

$$\begin{cases} \nabla \times \nabla \times \mathbf{u} + \mu \mathbf{u} = \mathbf{f}, & \Omega \\ \mathbf{u} \times \mathbf{n} = 0, & \partial\Omega \end{cases} \quad (12.18)$$

where  $\mathbf{f} \in \mathbf{L}^2(\Omega)$ ,  $\mu \in L^\infty(\Omega)$  and there exists  $\mu_0 > 0$  such that  $\mu \geq \mu_0$  almost everywhere. We take the Hilbert space  $V := \mathbf{H}_0(\text{curl}, \Omega)$ , in which case the variational formulation corresponding to (12.18) writes

Find  $\mathbf{u} \in V$  such that

$$a(\mathbf{u}, \mathbf{v}) = l(\mathbf{v}) \quad \forall \mathbf{v} \in V \quad (12.19)$$

where

$$\begin{cases} a(\mathbf{u}, \mathbf{v}) := \int_{\Omega} \nabla \times \mathbf{u} \cdot \nabla \times \mathbf{v} + \int_{\Omega} \mu \mathbf{u} \cdot \mathbf{v}, & \forall \mathbf{u}, \mathbf{v} \in V \\ l(\mathbf{v}) := \int_{\Omega} \mathbf{v} \cdot \mathbf{f}, & \forall \mathbf{v} \in V \end{cases} \quad (12.20)$$

We recall that in  $\mathbf{H}_0(\text{curl}, \Omega)$ , the bilinear form  $a$  is equivalent to the inner product and is therefor continuous and coercive. Hence, our abstract theory applies and there exists a unique solution to the problem (12.19).

 **$H(\text{div}, \Omega)$ -elliptic problem**

Let  $\Omega \subset \mathbb{R}^d$  be an open Lipschitz bounded set, and we look for the solution of the following problem

$$\begin{cases} -\nabla \nabla \cdot \mathbf{u} + \mu \mathbf{u} = \mathbf{f}, & \Omega \\ \mathbf{u} \times \mathbf{n} = 0, & \partial\Omega \end{cases} \quad (12.21)$$

where  $\mathbf{f} \in \mathbf{L}^2(\Omega)$ ,  $\mu \in L^\infty(\Omega)$  and there exists  $\mu_0 > 0$  such that  $\mu \geq \mu_0$  almost everywhere. We take the Hilbert space  $V := \mathbf{H}_0(\text{div}, \Omega)$ , in which case the variational formulation corresponding to (12.21) writes

Find  $\mathbf{u} \in V$  such that

$$a(\mathbf{u}, \mathbf{v}) = l(\mathbf{v}) \quad \forall \mathbf{v} \in V \quad (12.22)$$

where

$$\begin{cases} a(\mathbf{u}, \mathbf{v}) := \int_{\Omega} \nabla \cdot \mathbf{u} \nabla \cdot \mathbf{v} + \int_{\Omega} \mu \mathbf{u} \cdot \mathbf{v}, & \forall \mathbf{u}, \mathbf{v} \in V \\ l(\mathbf{v}) := \int_{\Omega} \mathbf{v} \cdot \mathbf{f}, & \forall \mathbf{v} \in V \end{cases} \quad (12.23)$$

We recall that in  $\mathbf{H}_0(\text{div}, \Omega)$ , the bilinear form  $a$  is equivalent to the inner product and is therefor continuous and coercive. Hence, our abstract theory applies and there exists a unique solution to the problem (12.22).

### 12.3 Saddle-point problems

In the sequel, we consider a special case of the problem (12.1).

Consider two Hilbert spaces  $V$  and  $W$ , two continuous bilinear forms  $a \in \mathcal{L}(V \times V, \mathbb{R})$  and  $b \in \mathcal{L}(V \times W, \mathbb{R})$  and two continuous linear forms  $l_V \in \mathcal{L}(V, \mathbb{R})$  and  $l_W \in \mathcal{L}(W, \mathbb{R})$ . We denote  $M_a$  and  $M_b$  the continuity constants for the bilinear forms  $a$  and  $b$  respectively. Then we define the abstract mixed variational problem as *Find*  $(u, p) \in V \times W$  such that

$$\begin{cases} a(u, v) + b(v, p) = l_V(v) & \forall v \in V \\ b(u, q) = l_W(q) & \forall q \in W \end{cases} \quad (12.24)$$

Many problems arising in CFD fit into this abstract framework, such as the Stokes equation. For saddle point problems the Lax-Milgram framework cannot be applied. The alternative solution is then to use the **inf-sup** conditions, known in this case as Banach-Nečas-Babuška (BNB) theorem. The link with the previous section is achieved by using the bilinear form  $c \in \mathcal{L}(X \times X, \mathbb{R})$

$$c((u, p), (v, q)) := a(u, v) + b(v, p) + b(u, q) \quad (12.25)$$

and the linear form  $l_X \in \mathcal{L}(X, \mathbb{R})$

$$l_X(v, q) := l_V(v) + l_W(q) \quad (12.26)$$

with  $X := V \times W$  endowed with the norm  $\|(u, p)\|_X := \|u\|_V + \|p\|_W$ .

Let us introduce the operators  $A : V \rightarrow V'$  and  $B : V \rightarrow W'$  such that

$$\langle Au, v \rangle_{V', V} := a(u, v) \quad \forall (u, v) \in V \times V \quad (12.27)$$

and

$$\langle Bu, p \rangle_{W', V} := b(u, p) \quad \forall (u, p) \in V \times W \quad (12.28)$$

Since all Hilbert spaces are reflexive Banach spaces, we have  $W'' = W$ . Hence we can define the following operator  $B^T : W \rightarrow V'$  such that

$$\langle B^T p, u \rangle_{V', W} := b(u, p) \quad \forall (u, p) \in V \times W \quad (12.29)$$

Therefore, the problem (12.24) is equivalent to *Find*  $(u, p) \in V \times W$  such that

$$\begin{cases} Au + B^T p = l_V \\ Bu = l_W \end{cases} \quad (12.30)$$

Now, let us introduce the nullspace of  $B$

$$\text{Ker } B := \{v \in V, \forall q \in W \quad b(v, q) = 0\} \quad (12.31)$$

The following theorem gives shows under which conditions the saddle problem (12.24) has a solution.

**Theorem 12.3.1** The variational problem (12.24) admits a unique solution if and only if

1) there exists  $\alpha > 0$ , such that

$$\begin{cases} \inf_{u \in \text{Ker } B} \sup_{v \in \text{Ker } B} \frac{a(u, v)}{\|u\|_V \|v\|_V} \geq \alpha \\ \forall v \in \text{Ker } B, \quad (\forall u \in \text{Ker } B, a(u, v) = 0) \Rightarrow (v = 0) \end{cases} \quad (12.32)$$

2) The Babuska-Brezzi, or inf-sup condition, is verified: there exists  $\beta > 0$  such that

$$\inf_{q \in W} \sup_{v \in V} \frac{b(v, q)}{\|v\|_V \|q\|_W} \geq \beta \quad (12.33)$$

In addition, the following a priori estimates hold

$$\begin{cases} \|u\|_V \leq \frac{1}{\alpha} \|l_V\|_{V'} + \frac{1}{\beta} \left(1 + \frac{M_a}{\alpha}\right) \|l_W\|_{W'} \\ \|p\|_W \leq \frac{1}{\beta} \left(1 + \frac{M_a}{\alpha}\right) \|l_V\|_{V'} + \frac{M_a}{\beta^2} \left(1 + \frac{M_a}{\alpha}\right) \|l_W\|_{W'} \end{cases} \quad (12.34)$$

A special case is when the bilinear form  $a$  is coercive. In this case, the first conditions can be replaced by a coercivity on  $\text{Ker } B$ .

**Theorem 12.3.2** Let  $V$  and  $W$  be Hilbert space. Assume  $a$  is a continuous bilinear form on  $V \times V$  and that  $b$  is a continuous linear form on  $V \times W$ , that  $l_V$  and  $l_W$  are continuous linear forms on  $V$  and  $W$  respectively and that the following two hypotheses are verified

1)  $a$  is coercive on  $K = \{v \in V \mid b(q, v) = 0, \forall q \in W\}$ , i.e. there exists  $\alpha > 0$  such that

$$a(v, v) \geq \alpha \|v\|_V^2 \quad \forall v \in K.$$

2) The Babuska-Brezzi, or inf-sup condition, is verified: there exists  $\beta > 0$  such that

$$\inf_{q \in W} \sup_{v \in V} \frac{b(v, q)}{\|q\|_W \|v\|_V} \geq \beta.$$

Then the variational problem admits a unique solution and the solution satisfies the a priori estimate

$$\begin{cases} \|u\|_V \leq \frac{1}{\alpha} \|l_V\|_{V'} + \frac{1}{\beta} \left(1 + \frac{M_a}{\alpha}\right) \|l_W\|_{W'} \\ \|p\|_W \leq \frac{1}{\beta} \left(1 + \frac{M_a}{\alpha}\right) \|l_V\|_{V'} + \frac{M_a}{\beta^2} \left(1 + \frac{M_a}{\alpha}\right) \|l_W\|_{W'} \end{cases} \quad (12.35)$$

The inf-sup conditions plays an essential role, as it is only satisfied if the spaces  $V$  and  $W$  are compatible in some sense. This condition being satisfied at the discrete level with a constant  $\beta$  that does not depend on the mesh size being essential for a well behaved Finite Element method. It can be written equivalently

$$\beta \|q\|_W \leq \sup_{v \in V} \frac{b(v, q)}{\|v\|_V} \quad \forall q \in W. \quad (12.36)$$

And often, a simple way to verify it is, given any  $q \in W$ , to find a specific  $v = v(q)$  depending on  $q$  such that

$$\beta \|q\|_W \leq \frac{b(v(q), q)}{\|v(q)\|_V} \leq \sup_{v \in V} \frac{b(v, q)}{\|v\|_V}$$

with a constant  $\beta$  independent of  $w$ .

### 12.3.1 Examples

#### First mixed formulation of the Poisson problem

Let  $\Omega \subset \mathbb{R}^3$  and consider the Poisson problem

$$\begin{cases} -\Delta p = f & , \Omega \\ p = 0 & , \partial\Omega \end{cases} \quad (12.37)$$

Using that  $\Delta p = \nabla \cdot \nabla p$ , we set  $\mathbf{u} = \nabla p$ , then the Poisson equation (12.37) can be written equivalently

$$\mathbf{u} = -\nabla p, \quad \nabla \cdot \mathbf{u} = f.$$

Instead of having one unknown, we now have two, along with the above two equations. In order to get a mixed variational formulation, we first take the dot product of the first one by  $\mathbf{v}$  and integrate

by parts

$$\int_{\Omega} \mathbf{u} \cdot \mathbf{v} \, dx - \int_{\Omega} p \nabla \cdot \mathbf{v} \, dx + \int_{\partial\Omega} p \mathbf{v} \cdot \mathbf{n} \, d\sigma = \int_{\Omega} \mathbf{u} \cdot \mathbf{v} \, dx - \int_{\Omega} p \nabla \cdot \mathbf{v} \, dx = 0,$$

using  $p = 0$  as a natural boundary condition. Then multiplying the second equation by  $q$  and integrating yields

$$\int_{\Omega} \nabla \cdot \mathbf{u} q \, dx = \int_{\Omega} f q \, dx.$$

No integration by parts is necessary here. And we thus get the following mixed variational formulation:

Find  $(\mathbf{u}, p) \in H(\text{div}, \Omega) \times L^2(\Omega)$  such that

$$\begin{cases} \int_{\Omega} \mathbf{u} \cdot \mathbf{v} \, dx - \int_{\Omega} p \nabla \cdot \mathbf{v} \, dx = 0, & \forall \mathbf{v} \in H(\text{div}, \Omega) \\ - \int_{\Omega} \nabla \cdot \mathbf{u} q \, dx = - \int_{\Omega} f q \, dx, & \forall q \in L^2(\Omega) \end{cases} \quad (12.38)$$

### Second mixed formulation of the Poisson problem

Here, we get an alternative formulation by not integrating by parts, the mixed term in the first formulation but in the second. The first formulation simply becomes

$$\int_{\Omega} \mathbf{u} \cdot \mathbf{v} \, dx + \int_{\Omega} \nabla p \cdot \mathbf{v} \, dx = 0,$$

and the second, removing immediately the boundary term due to the essential boundary condition  $q = 0$

$$\int_{\Omega} \nabla \cdot \mathbf{u} q \, dx = - \int_{\Omega} \mathbf{u} \cdot \nabla q \, dx = \int_{\Omega} f q \, dx,$$

which leads to the variational formulation

Find  $(\mathbf{u}, p) \in L^2(\Omega)^3 \times H_0^1(\Omega)$  such that

$$\begin{cases} \int_{\Omega} \mathbf{u} \cdot \mathbf{v} \, dx + \int_{\Omega} \nabla p \cdot \mathbf{v} \, dx = 0, & \forall \mathbf{v} \in L^2(\Omega)^3 \\ \int_{\Omega} \mathbf{u} \cdot \nabla q \, dx = - \int_{\Omega} f q \, dx, & \forall q \in H_0^1(\Omega) \end{cases} \quad (12.39)$$

Note that this formulation actually contains the classical variational formulation for the Poisson equation. Indeed for  $q \in H_0^1(\Omega)$ ,  $\nabla q \in L^2(\Omega)^3$  can be used as a test function in the first equation. And plugging this into the second we get

$$\int_{\Omega} \nabla p \cdot \nabla q \, dx = \int_{\Omega} f q \, dx, \quad \forall q \in H_0^1(\Omega).$$

### First mixed formulation of the Stokes problem

We consider now the Stokes problem for the steady-state modelling of an incompressible fluid

$$\begin{cases} -\nabla^2 \mathbf{u} + \nabla p = \mathbf{f} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = 0 & \text{on } \partial\Omega, \end{cases} \quad (12.40)$$

For the variational formulation, we take the dot product of the first equation with  $v$  and integrate over the whole domain

$$\int_{\Omega} (-\Delta \mathbf{u} + \nabla p) \cdot \mathbf{v} \, dx = \int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} \, dx + \int_{\Omega} \nabla p \cdot \mathbf{v} \, dx = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, dx$$

The integration by parts is performed component by component. We impose the essential boundary condition  $\mathbf{v} = 0$  on  $\partial\Omega$ , and we denote by

$$\int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} \, d\mathbf{x} = \sum_{i=1}^3 \int_{\Omega} \nabla u_i \cdot \nabla v_i \, d\mathbf{x} = \sum_{i,j=1}^3 \int_{\Omega} \partial_j u_i \partial_j v_i \, d\mathbf{x}.$$

We now need to deal with the constraint  $\nabla \cdot \mathbf{u} = 0$ . The theoretical framework for saddle point problems requires that the corresponding bilinear form is the same as the second one appearing in the first part of the variational formulation. To this aim we multiply  $\nabla \cdot \mathbf{u} = 0$  by a scalar test function (which will be associated to  $p$ ) and integrate on the full domain, with an integration by parts in order to get the same bilinear form as in the first equation

$$\int_{\Omega} \nabla \cdot \mathbf{u} q \, d\mathbf{x} = - \int_{\Omega} \mathbf{u} \cdot \nabla q \, d\mathbf{x} = 0,$$

using that  $q = 0$  on the boundary as an essential boundary condition. We finally obtain the mixed variational formulation:

Find  $(\mathbf{u}, p) \in H_0^1(\Omega)^3 \times H_0^1(\Omega)$  such that

$$\begin{cases} \int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} \, d\mathbf{x} + \int_{\Omega} \nabla p \cdot \mathbf{v} \, d\mathbf{x} = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, d\mathbf{x}, & \forall \mathbf{v} \in H_0^1(\Omega)^3 \\ \int_{\Omega} \mathbf{u} \cdot \nabla q \, d\mathbf{x} = 0, & \forall q \in H_0^1(\Omega) \end{cases} \quad (12.41)$$

### Second mixed formulation of the Stokes problem

Another possibility to obtain a well posed variational formulation, is to integrate by parts the  $\int_{\Omega} \nabla p \cdot \mathbf{v} \, d\mathbf{x}$  term in the first formulation:

$$\int_{\Omega} \nabla p \cdot \mathbf{v} \, d\mathbf{x} = - \int_{\Omega} p \nabla \cdot \mathbf{v} \, d\mathbf{x} + \int_{\partial\Omega} p \mathbf{v} \cdot \mathbf{n} \, d\sigma = - \int_{\Omega} p \nabla \cdot \mathbf{v} \, d\mathbf{x},$$

using here  $p = 0$  as a natural boundary condition. Note that in the other variational formulation the same boundary condition was essential. In this case, for the second variational formulation, we just multiply  $\nabla \cdot \mathbf{u} = 0$  by  $q$  and integrate. No integration by parts is needed in this case.

$$\int_{\Omega} \nabla \cdot \mathbf{u} q \, d\mathbf{x} = 0.$$

This then leads to the following variational formulation:

Find  $(\mathbf{u}, p) \in H^1(\Omega)^3 \times L^2(\Omega)$  such that

$$\begin{cases} \int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} \, d\mathbf{x} - \int_{\Omega} p \nabla \cdot \mathbf{v} \, d\mathbf{x} = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, d\mathbf{x}, & \forall \mathbf{v} \in H^1(\Omega)^3 \\ \int_{\Omega} \nabla \cdot \mathbf{u} q \, d\mathbf{x} = 0, & \forall q \in L^2(\Omega) \end{cases} \quad (12.42)$$

### 12.3.2 Galerkin approximation

Let us now come to the Galerkin discretisation. The principle is simply to construct finite dimensional subspaces  $W_h \subset W$  and  $V_h \subset V$  and to write the variational formulation (12.24) replacing  $W$  by  $W_h$  and  $V$  by  $V_h$ . The variational formulations are the same as in the continuous case, like for conforming finite elements. This automatically yields the consistency of the discrete formulation. In order to get the stability property needed for convergence, we need that the coercivity constant  $\alpha$  and the inf-sup constant  $\beta$  are independent of  $h$ .

Because  $V_h \subset V$  the coercivity property is automatically verified in the discrete case, with a coercivity constant that is the same as in the continuous case and hence does not depend on the discretisation parameter  $h$ .

Here, however, there is an additional difficulty, linked to the inf-sup conditions, which is completely dependent on the two spaces  $V_h$  and  $W_h$ . By far not any conforming approximation of the two spaces will verify the discrete inf-sup condition with a constant  $\beta$  that is independent on  $h$ . Finding compatible discrete spaces for a given mixed variational formulation, has been an active area of research.

The variational problem for the Galerkin approximation is *Find*  $(u_h, p_h) \in V_h \times W_h$  such that

$$\begin{cases} a(u_h, v_h) + b(v_h, p_h) = l_V(v_h) & \forall v_h \in V_h \\ b(u_h, q_h) = l_W(q_h) & \forall q_h \in W_h \end{cases} \quad (12.43)$$

Let us introduce the operator  $B_h : V_h \rightarrow W_h'$  such that

$$\langle B_h u_h, p_h \rangle_{W_h', V_h} := b(u_h, p_h) \quad \forall (u_h, p_h) \in V_h \times W_h \quad (12.44)$$

and its nullspace

$$\text{Ker } B_h := \{v_h \in V_h, \forall q_h \in W_h \quad b(v_h, q_h) = 0\} \quad (12.45)$$

The following proposition states the conditions under which the Galerkin approximation of the problem (12.43) admits a solution

**Proposition 12.3.3** The variational problem (12.43) admits a unique solution if and only if

- 1) there exists  $\alpha_h > 0$ , such that

$$\inf_{u_h \in \text{Ker } B_h} \sup_{v_h \in \text{Ker } B_h} \frac{a(u_h, v_h)}{\|u_h\|_V \|v_h\|_V} \geq \alpha_h \quad (12.46)$$

- 2) there exists  $\beta_h > 0$  such that

$$\inf_{q_h \in W_h} \sup_{v_h \in V_h} \frac{b(u_h, q_h)}{\|v_h\|_V \|q_h\|_W} \geq \beta_h \quad (12.47)$$

**R** The second condition is equivalent to assuming  $B_h$  is surjective.

Finally, we state the following lemma which is equivalent to Cea's lemma.

**Lemma 13** Under the assumptions of theorem 12.3.3, we have

- a) if  $\text{Ker } B_h \subset \text{Ker } B$ ,

$$\begin{cases} \|u - u_h\|_V \leq \left(1 + \frac{M_a}{\alpha_h}\right) \left(1 + \frac{M_b}{\beta_h}\right) \inf_{v_h \in V_h} \|u - v_h\|_V \\ \|p - p_h\|_W \leq \frac{M_a}{\beta_h} \left(1 + \frac{M_a}{\alpha_h}\right) \left(1 + \frac{M_b}{\beta_h}\right) \inf_{v_h \in V_h} \|u - v_h\|_V \\ \quad + \left(1 + \frac{M_b}{\beta_h}\right) \inf_{q_h \in W_h} \|p - q_h\|_W \end{cases} \quad (12.48)$$

- b) otherwise,

$$\begin{cases} \|u - u_h\|_V \leq \left(1 + \frac{M_a}{\alpha_h}\right) \left(1 + \frac{M_b}{\beta_h}\right) \inf_{v_h \in V_h} \|u - v_h\|_V \\ \quad + \frac{M_b}{\alpha_h} \inf_{q_h \in W_h} \|p - q_h\|_W \\ \|p - p_h\|_W \leq \frac{M_a}{\beta_h} \left(1 + \frac{M_a}{\alpha_h}\right) \left(1 + \frac{M_b}{\beta_h}\right) \inf_{v_h \in V_h} \|u - v_h\|_V \\ \quad + \left(1 + \frac{M_b}{\beta_h} + \frac{M_a}{\alpha_h} \frac{M_b}{\beta_h}\right) \inf_{q_h \in W_h} \|p - q_h\|_W \end{cases} \quad (12.49)$$

### 12.3.3 Examples

#### Matrix form of the first mixed formulation of the Poisson problem

Let  $V_h$  and  $W_h$  be subspaces of finite dimensions of  $\mathbf{H}(\text{div}, \Omega)$  and  $L^2(\Omega)$  respectively, leading to a stable discretization of the variational problem (12.38). We shall assume that

$$V_h = \text{span}\{\Psi_i, 1 \leq i \leq N_{V_h}\}$$

and

$$W_h = \text{span}\{\phi_i, 1 \leq i \leq N_{W_h}\}$$

where  $N_{V_h}$  and  $N_{W_h}$  is the dimension of  $V_h$  and  $W_h$  respectively. For  $\mathbf{u}_h \in V_h$  and  $p_h \in W_h$ , we can write

$$\mathbf{u}_h = \sum_{j=1}^{N_{V_h}} u_j \Psi_j, \quad p_h = \sum_{j=1}^{N_{W_h}} p_j \phi_j$$

By taking  $\mathbf{v} = \Psi_i$  in the first equation of the variational formulation, we get

$$\sum_{j=1}^{N_{V_h}} u_j \int_{\Omega} \Psi_j \cdot \Psi_i \, d\mathbf{x} - \sum_{j=1}^{N_{W_h}} p_j \int_{\Omega} \phi_j \nabla \cdot \Psi_i \, d\mathbf{x} = 0, \quad \forall 1 \leq i \leq N_{V_h}$$

By taking  $q = \phi_i$  in the second equation of the variational formulation, we get

$$\sum_{j=1}^{N_{\text{div}}} u_j \int_{\Omega} \nabla \cdot \Psi_j \phi_i \, d\mathbf{x} = \int_{\Omega} f \phi_i \, d\mathbf{x}, \quad \forall 1 \leq i \leq N_{W_h}$$

Find  $(U, P) \in \mathbb{R}^{N_{V_h}} \times \mathbb{R}^{N_{W_h}}$  such that

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix} \quad (12.50)$$

where the matrices  $A$  and  $B$  are given by

$$\begin{aligned} A_{i,j} &:= \int_{\Omega} \Psi_j \cdot \Psi_i \, d\mathbf{x}, \quad 1 \leq i, j \leq N_{V_h} \\ B_{i,j} &:= - \int_{\Omega} \phi_j \nabla \cdot \Psi_i \, d\mathbf{x}, \quad 1 \leq j \leq N_{V_h}, \quad 1 \leq i \leq N_{W_h} \\ F_i &:= - \int_{\Omega} f \phi_i \, d\mathbf{x}, \quad 1 \leq i \leq N_{W_h} \end{aligned}$$

#### Matrix form for the second mixed formulation of the Poisson problem

Let  $V_h$  and  $W_h$  be subspaces of finite dimensions of  $L^2(\Omega)^3$  and  $H_0^1(\Omega)$  respectively, leading to a stable discretization of the variational problem (12.39). We shall assume that

$$V_h = \text{span}\{\Psi_i, 1 \leq i \leq N_{V_h}\}$$

and

$$W_h = \text{span}\{\phi_i, 1 \leq i \leq N_{W_h}\}$$

where  $N_{V_h}$  and  $N_{W_h}$  is the dimension of  $V_h$  and  $W_h$  respectively. Following the same approach as before, we get the matrix form of our variational formulation



Find  $(U, P) \in \mathbb{R}^{N_{V_h}} \times \mathbb{R}^{N_{W_h}}$  such that

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix} \quad (12.51)$$

where the matrices  $A$  and  $B$  are given by

$$\begin{aligned} A_{i,j} &:= \int_{\Omega} \boldsymbol{\Psi}_j \cdot \boldsymbol{\Psi}_i \, d\mathbf{x}, \quad 1 \leq i, j \leq N_{V_h} \\ B_{i,j} &:= \int_{\Omega} \nabla \phi_j \cdot \boldsymbol{\Psi}_i \, d\mathbf{x}, \quad 1 \leq j \leq N_{V_h}, \quad 1 \leq i \leq N_{W_h} \\ F_i &:= - \int_{\Omega} f \phi_i \, d\mathbf{x}, \quad 1 \leq i \leq N_{W_h} \end{aligned}$$

### First mixed formulation of the Stokes problem

Let  $V_h$  and  $W_h$  be subspaces of finite dimensions of  $H_0^1(\Omega)^3$  and  $H_0^1(\Omega)$  respectively, leading to a stable discretization of the variational problem (12.41). We shall assume that

$$V_h = \text{span}\{\boldsymbol{\Psi}_i, 1 \leq i \leq N_{V_h}\}$$

and

$$W_h = \text{span}\{\phi_i, 1 \leq i \leq N_{W_h}\}$$

where  $N_{V_h}$  and  $N_{W_h}$  is the dimension of  $V_h$  and  $W_h$  respectively. Following the same approach as before, we get the matrix form of our variational formulation

Find  $(U, P) \in \mathbb{R}^{N_{V_h}} \times \mathbb{R}^{N_{W_h}}$  such that

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix} \quad (12.52)$$

where the matrices  $A$  and  $B$  are given by

$$\begin{aligned} A_{i,j} &:= \int_{\Omega} \boldsymbol{\Psi}_j : \boldsymbol{\Psi}_i \, d\mathbf{x}, \quad 1 \leq i, j \leq N_{V_h} \\ B_{i,j} &:= \int_{\Omega} \nabla \phi_j \cdot \boldsymbol{\Psi}_i \, d\mathbf{x}, \quad 1 \leq j \leq N_{V_h}, \quad 1 \leq i \leq N_{W_h} \\ F_i &:= \int_{\Omega} \mathbf{f} \cdot \boldsymbol{\Psi}_i \, d\mathbf{x}, \quad 1 \leq i \leq N_{W_h} \end{aligned}$$

### Second mixed formulation of the Stokes problem

Let  $V_h$  and  $W_h$  be subspaces of finite dimensions of  $H^1(\Omega)^3$  and  $L^2(\Omega)$  respectively, leading to a stable discretization of the variational problem (12.42). We shall assume that

$$V_h = \text{span}\{\boldsymbol{\Psi}_i, 1 \leq i \leq N_{V_h}\}$$

and

$$W_h = \text{span}\{\phi_i, 1 \leq i \leq N_{W_h}\}$$

where  $N_{V_h}$  and  $N_{W_h}$  is the dimension of  $V_h$  and  $W_h$  respectively. Following the same approach as before, we get the matrix form of our variational formulation

Find  $(U, P) \in \mathbb{R}^{N_{V_h}} \times \mathbb{R}^{N_{W_h}}$  such that

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix} \quad (12.53)$$

where the matrices  $A$  and  $B$  are given by

$$A_{i,j} := \int_{\Omega} \boldsymbol{\Psi}_j : \boldsymbol{\Psi}_i \, d\mathbf{x}, \quad 1 \leq i, j \leq N_{V_h}$$

$$B_{i,j} := - \int_{\Omega} \phi_j \nabla \cdot \boldsymbol{\Psi}_i \, d\mathbf{x}, \quad 1 \leq j \leq N_{V_h}, \quad 1 \leq i \leq N_{W_h}$$

$$F_i := \int_{\Omega} \mathbf{f} \cdot \boldsymbol{\Psi}_i \, d\mathbf{x}, \quad 1 \leq i \leq N_{W_h}$$

## 12.4 Problems

TODO

## 13. Historical Notes

TODO

# IV Isogeometric Analysis

<b>14</b>	<b>Isogeometric Finite Elements .....</b>	<b>118</b>
14.1	Sobolev estimations under h-refinement	
14.2	Galerkin approximation	
14.3	Exact DeRham sequences	
14.4	Problems	
<b>15</b>	<b>Historical Notes .....</b>	<b>139</b>



# 14. Isogeometric Finite Elements

## 14.1 Sobolev estimations under h-refinement

### 14.1.1 Approximation properties without mapping

We consider a computational domain  $\hat{\Omega} := (0, 1)^d$ , with  $d \geq 1$ . Given  $d$  knot vectors for each direction  $T^{(\alpha)} := \{t_0^{(\alpha)}, \dots, t_{n_\alpha+p_\alpha+1}^{(\alpha)}\}$ . For  $\alpha \in \{1, \dots, d\}$ , we denote  $I_i^{(\alpha)} := (t_i^{(\alpha)}, t_{i+1}^{(\alpha)})$  and by construction we have  $t_{p_\alpha}^{(\alpha)} = 0$  and  $t_{n_\alpha}^{(\alpha)} = 1$ . We shall denote  $I := (0, 1)$ , *i.e.*  $\hat{\Omega} = I^d$ . The  $d$  knot vectors will define a natural triangulation of  $\hat{\Omega}$ , which we denote  $\mathcal{Q}_h(\hat{\Omega})$  (the subscript  $h$  will be made clear later). An element of  $Q \in \mathcal{Q}_h(\hat{\Omega})$  is of the form  $\bigotimes_{1 \leq \alpha \leq d} I_{i_\alpha}^{(\alpha)}$ , for some multi-index  $\mathbf{i} := (i_1, \dots, i_d)$ . We will denote the element size of  $Q \in \mathcal{Q}_h(\hat{\Omega})$  by  $h_Q := \text{diam}(Q)$ , moreover, we shall define  $h_{Q,\alpha}$  for  $1 \leq \alpha \leq d$  as the length of  $I_{i_\alpha}^{(\alpha)}$ , while  $h := \max_{Q \in \mathcal{Q}_h(\hat{\Omega})} h_Q$  will represent the global mesh size.

We shall also assume our mesh to be **locally quasi-uniform**, meaning, there exists a constant  $\theta \geq 1$  such that

$$\frac{1}{\theta} \leq \frac{h_{I_{i_\alpha}^{(\alpha)}, \alpha}}{h_{I_{i_{\alpha+1}}^{(\alpha)}, \alpha}} \leq \theta, \quad \forall 1 \leq \alpha \leq d \quad (14.1)$$

For every element  $Q$  we define its extension as the union of supports of non-vanishing B-Splines on  $Q$ , which we will denote by  $\tilde{Q}$ , (see figure ??). Finally, the  $\alpha$ -coordinate will be denoted by  $\eta_\alpha$  for all  $1 \leq \alpha \leq d$ .

We recall that the Schoenberg space of degree  $p$  and associated to the knot sequence  $T$  is denoted by  $\mathcal{S}_p(T)$ . Another equivalent notation would be  $\mathcal{S}_\mu^p(T^*)$  (or simply  $\mathcal{S}_\mu^p$  for the sake of simplicity), where as usual  $T^*$  denotes the set of breakpoints associated to  $T$  and  $\mu$  is a sequence of

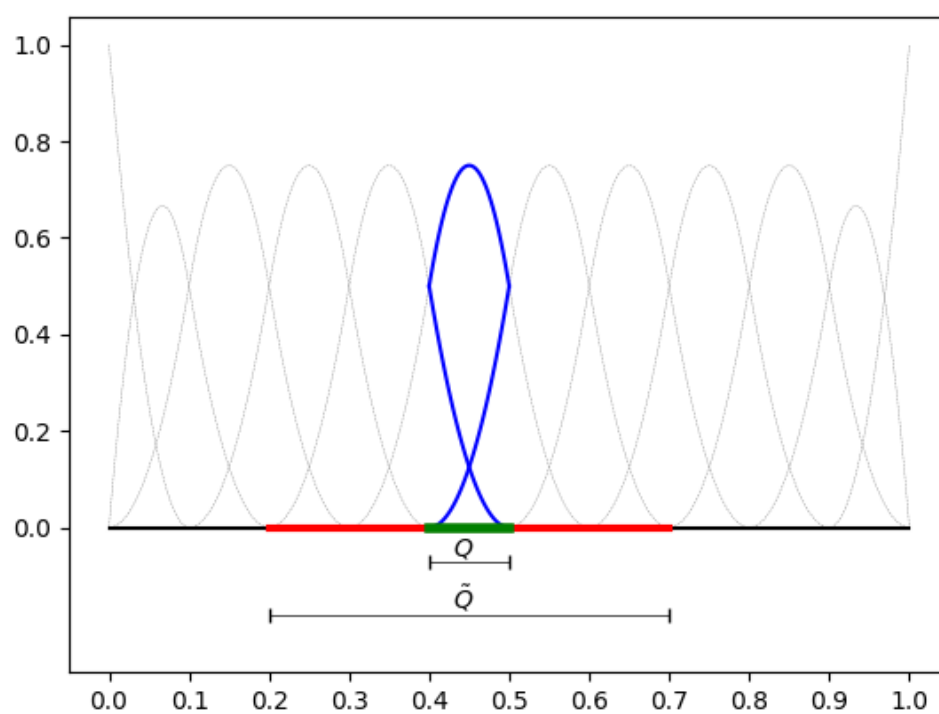


Figure 14.1: An element  $Q$  and its extension  $\tilde{Q}$  for an open knot sequence, with quadratic B-Splines.

multiplicities of each breakpoint. More generally, in higher dimensions, the Schoenberg space is defined by tensor product

$$V_h := \bigotimes_{1 \leq \alpha \leq d} \mathcal{S}_{\mu_\alpha}^{p_\alpha} \quad (14.2)$$

Finally, we shall denote by  $\Pi_h$  a quasi-interpolant on  $V_h$ .

Let us introduce the following space of functions, which will be needed to define the bent Sobolev space:

$$\mathcal{D}_T^s(I) := \{f \in L^2(I); D_-^k f(t_i) = D_+^k f(t_i), \quad \forall 0 \leq k \leq \min(s-1, \mu_i), \forall 1 \leq i \leq n-1\} \quad (14.3)$$

We define the **bent Sobolev space** on  $I$  as

$$\mathcal{H}^s(I) = \{f \in L^2(I), f|_Q \in H^s(Q), \forall Q \in \mathcal{Q}_h(I) \text{ and } f \in \mathcal{D}_T^s\} \quad (14.4)$$

The bent Sobolev spaces are endowed with the broken norm and seminorms

$$\|f\|_{\mathcal{H}^s(I)}^2 := \sum_{0 \leq j \leq s} |f|_{\mathcal{H}^j(I)}^2 \quad (14.5)$$

where

$$|f|_{\mathcal{H}^j(I)}^2 := \sum_{I_i \in \mathcal{Q}_h(I)} |f|_{H^j(I_i)}^2 \quad (14.6)$$

In higher dimensions, the bent Sobolev spaces are defined by tensor product, for a multi-index  $\mathbf{s} = (s_1, \dots, s_d)$ , as follows

$$\mathcal{H}^{\mathbf{s}}(\hat{\Omega}) = \bigotimes_{1 \leq \alpha \leq d} \mathcal{H}^{s_\alpha}(I) \quad (14.7)$$

which are endowed by the tensor-product norm and seminorms.

For any sufficiently regular function  $f : \hat{\Omega} \rightarrow \mathbb{R}$ , we shall define the partial derivative operator of order  $\mathbf{r}$  (a multi-index)

$$D^{\mathbf{r}} f := \frac{\partial^{r_1} \dots \partial^{r_d}}{\partial \eta_1^{r_1} \dots \partial \eta_d^{r_d}} f \quad (14.8)$$

Finally, for any union  $A \subset \hat{\Omega}$  of elements of the triangulation  $\mathcal{Q}_h(\hat{\Omega})$ , we will define

$$\|f\|_{L_h^2(A)}^2 := \sum_{Q \in \mathcal{Q}_h(\hat{\Omega}) \text{ and } Q \subset A} \|f\|_{L^2(Q)}^2 \quad (14.9)$$

The following proposition gives a local anisotropic estimation of the best approximation.

**Proposition 14.1.1 — Anisotropic estimation in 2d.** If  $0 \leq \mathbf{r} \leq \mathbf{s} \leq \mathbf{p} + 1$ , there exists a constant  $C(\mathbf{p}, \boldsymbol{\theta})$  such that for all elements  $Q \in \mathcal{Q}_h(\hat{\Omega})$ , we have

$$\|D^{\mathbf{r}}(f - \Pi_f)\|_{L^2(Q)} \leq C \left( h_{Q,1}^{s_1-r_1} \|D^{(s_1,r_2)} f\|_{L_h^2(\tilde{Q})} + h_{Q,2}^{s_2-r_2} \|D^{(r_1,s_2)} f\|_{L_h^2(\tilde{Q})} \right) \quad (14.10)$$

for all  $f \in \mathcal{H}^{\mathbf{s}}(\hat{\Omega}) \cap \mathcal{H}^{\mathbf{r}}(\hat{\Omega})$ .



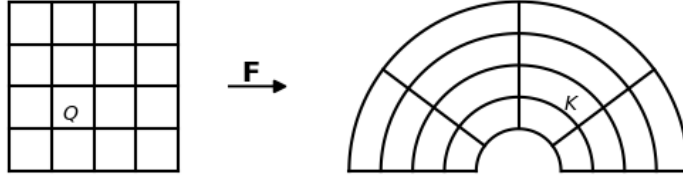


Figure 14.2: The triangulation  $\mathcal{Q}_h(\hat{\Omega})$  in the logical domain and its image  $\mathcal{Q}_h(\Omega)$ , under the mapping  $\mathbf{F}$ . An element  $Q \in \mathcal{Q}_h(\hat{\Omega})$  is mapped onto  $K \in \mathcal{Q}_h(\Omega)$ .

### 14.1.2 Approximation properties with mapping

We now turn to the general case, where the domain  $\Omega := \mathbf{F}(\hat{\Omega})$  with  $\mathbf{F} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a one to one map. We shall use the results from the previous section on  $\hat{\Omega}$ , and then assume that  $\hat{\Omega} := (0, 1)^d$ . The  $\alpha$ -coordinate will be denoted, in this case, by  $x_\alpha$  for all  $1 \leq \alpha \leq d$ .  $\Omega$  will be subdivided into elements  $K := \mathbf{F}(Q)$  where  $Q \in \mathcal{Q}_h(\hat{\Omega})$ , hence  $\mathcal{Q}_h(\Omega) := \mathbf{F}(\mathcal{Q}_h(\hat{\Omega}))$ . Moreover, every extension  $\tilde{K}$  of  $K$  is also of the form  $\mathbf{F}(\tilde{Q})$  with  $\tilde{K} = \mathbf{F}(\tilde{Q})$ .

In order to define the bent Sobolev space, we will first introduce the following broken norms and seminorms

$$\|f\|_{\mathcal{H}_{\mathbf{F}}^s(A)}^2 := \sum_{0 \leq j \leq s} |f|_{\mathcal{H}_{\mathbf{F}}^j(A)}^2 \quad (14.11)$$

where

$$|f|_{\mathcal{H}_{\mathbf{F}}^s(A)}^2 := \sum_{K \in \mathcal{Q}_h(\Omega) \text{ s.t. } K \subset A} \|D_{\mathbf{F}}^s f\|_{L^2(K)}^2 \quad (14.12)$$

Finally, we shall consider the *bent Sobolev space*  $\mathcal{H}_{\mathbf{F}}^s$  as the **closure** of  $C^\infty(\Omega)$  with respect to the norm  $\|\cdot\|_{\mathcal{H}_{\mathbf{F}}^s}$ . We can now state the approximation theorem on the bent Sobolev space, which is suitable for anisotropic estimates.

**Theorem 14.1.2** If  $0 \leq \mathbf{r} \leq \mathbf{s} \leq \mathbf{p} + 1$ , there exists a constant  $C(\mathbf{p}, \boldsymbol{\theta}, \mathbf{F})$  such that for all elements  $K \in \mathcal{Q}_h(\Omega)$ , we have

$$\|f - \Pi_h f\|_{\mathcal{H}_{\mathbf{F}}^{\mathbf{r}}(K)} \leq C \left( h_{K,1}^{s_1 - r_1} \|f\|_{\mathcal{H}_{\mathbf{F}}^{(s_1, r_2)}(\tilde{K})} + h_{K,2}^{s_2 - r_2} \|f\|_{\mathcal{H}_{\mathbf{F}}^{(r_1, s_2)}(\tilde{K})} \right) \quad (14.13)$$

for all  $f \in \mathcal{H}^{(s_1, r_2)}(\Omega) \cap \mathcal{H}^{(r_1, s_2)}(\Omega)$ .

**Corollary 14.1.3** If  $0 \leq r \leq s \leq \min(\mathbf{p}) + 1$ , there exists a constant  $C(\mathbf{p}, \theta, \mathbf{F})$  such that for all elements  $K \in \mathcal{Q}_h(\Omega)$ , we have

$$\|f - \Pi_h f\|_{H^r(K)} \leq Ch_K^{s-r} \|f\|_{H^s(\bar{K})} \quad (14.14)$$

and

$$\|f - \Pi_h f\|_{H^r(\Omega)} \leq Ch^{s-r} \|f\|_{H^s(\Omega)} \quad (14.15)$$

for all  $f \in H^s(\Omega)$ .

## 14.2 Galerkin approximation

As in the previous section, we shall distinguish between the two cases with and without mapping.

### 14.2.1 Case without mapping

We consider a first coarse mesh, which is given in this case by one element in each direction  $\hat{\Omega} = (0, 1)^d$ . Then we define a sequence of  $d$  knot vectors  $(\{T_N^{(\alpha)}, 1 \leq \alpha \leq d\})_{N \geq 0}$  such that  $\lim_{N \rightarrow \infty} h_N = 0$  where  $h_N$  is the diameter associated to the triangulation endowed by the knot vectors  $\{T_N^{(\alpha)}, 1 \leq \alpha \leq d\}$ .

**R** Creating such a sequence of knot vectors can be done through the insertion of knots, as presented in the CAD part of this lecture.

For every  $N$ , we have a Schoenberg space associated to the knot vectors  $\{T_N^{(\alpha)}, 1 \leq \alpha \leq d\}$ , which is denoted by  $V_h$ .

#### Matrix form for Poisson:

We consider the Poisson problem with Homogeneous Dirichlet boundary conditions, for which the variational problem writes: *Find  $u \in H_0^1(\hat{\Omega})$  such that*

$$\int_{\hat{\Omega}} \nabla u \cdot \nabla v \, d\boldsymbol{\eta} = \int_{\hat{\Omega}} f v \, d\boldsymbol{\eta}, \quad \forall v \in H_0^1(\hat{\Omega}) \quad (14.16)$$

One can create a subspace of  $H_0^1(\hat{\Omega})$  from the tensor product of Schoenberg spaces (14.2), by removing the first and last B-Splines, which are interpolatories (equal to 1) at the endpoints in every direction, which we will denote by  $V_h^0$ . From now on, we shall use multi-indices and introduce the tensor B-Splines functions

$$N_{\mathbf{i}}^{\mathbf{p}}(\boldsymbol{\eta}) := \prod_{1 \leq \alpha \leq d} N_{i_{\alpha}}^{p_{\alpha}}(\eta_{\alpha}) \quad (14.17)$$

where  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_d)$ ,  $\mathbf{p} = (p_1, \dots, p_d)$  and  $\mathbf{i} = (i_1, \dots, i_d)$ . Using these notations, we have

$$V_h = \text{span}\{N_{\mathbf{i}}^{\mathbf{p}}; 0 \leq \mathbf{i} \leq \mathbf{n}\} \quad (14.18)$$

while

$$V_h^0 = \text{span}\{N_{\mathbf{i}}^{\mathbf{p}}; 1 \leq \mathbf{i} \leq \mathbf{n} - 1\} \quad (14.19)$$

where  $\mathbf{n} = (n_1, \dots, n_d)$ .

The discrete variational formulation is: Find  $u \in V_h^0$  such that

$$\int_{\hat{\Omega}} \nabla u \cdot \nabla v \, d\boldsymbol{\eta} = \int_{\hat{\Omega}} f v \, d\boldsymbol{\eta}, \quad \forall v \in V_h^0 \quad (14.20)$$

Let  $u_h \in V_h^0$  such that  $u_h = \sum_j u_j N_{\boldsymbol{\mu},j}^p$ . Then the weak formulation for the Poisson equation writes

$$\sum_j u_j \int_{\hat{\Omega}} \nabla N_{\boldsymbol{\mu},j}^p \cdot \nabla N_{\boldsymbol{\mu},i}^p \, d\boldsymbol{\eta} = \int_{\hat{\Omega}} f N_{\boldsymbol{\mu},i}^p \, d\boldsymbol{\eta}, \quad \forall 1 \leq i \leq n-1$$

In practice, in order to store these terms, we will need to have a reordering of the indices so that we map the set of multi-indices  $\{1 \leq \mathbf{i} \leq \mathbf{n}-1\}$  onto a set of 1d indices  $\{1 \leq I \leq N\}$ . Using this reordering, we get the following linear system

$$MU = F$$

where  $U$  denotes the coefficients  $(u_J, 1 \leq J \leq N)$  and  $F$  is a vector consisting of the terms  $\int_{\hat{\Omega}} f N_{\boldsymbol{\mu},i}^p \, d\boldsymbol{\eta}$  for  $1 \leq i \leq \mathbf{n}-1$ . Finally, the matrix  $M$  is called the stiffness matrix and its entries are defined as

$$M_{IJ} = \int_{\hat{\Omega}} \nabla N_{\boldsymbol{\mu},j}^p \cdot \nabla N_{\boldsymbol{\mu},i}^p \, d\boldsymbol{\eta}$$

**Exercise 14.1** We assume in this exercise  $d = 2$ . Show that solving the associated linear system to the Poisson problem is equivalent to Find  $X \in \mathbb{R}^{n_1-1} \times \mathbb{R}^{n_2-1}$ , such that

$$S_1 X M_2 + M_1 X S_2 = F \quad (14.21)$$

where the vector of unknowns (and right hand side) are *viewed* as matrices in  $X \in \mathbb{R}^{n_1-1} \times \mathbb{R}^{n_2-1}$ .

#### Matrix form for the Biharmonic problem:

We consider the Biharmonic problem (12.15), for which the variational formulation is given by Find  $u \in V$  such that

$$a(u, v) = l(v) \quad \forall v \in V \quad (14.22)$$

with  $V := H_0^2(\Omega)$  and

$$a(u, v) := \int_{\Omega} \Delta u \Delta v \, dx, \quad \langle L, v \rangle = \int_{\Omega} f v \, dx$$

Because of the Sobolev embedding, functions in  $V$  can be seen as  $C^1(\Omega)$ . Therefore, we have to create the Schoenberg spaces in such a way the global regularity is at least  $C^1$ . This leads to a condition on the multiplicity of the interior knots and is ensured if we take  $\mathbf{p} \geq \boldsymbol{\mu} + 1$ , with  $\mathbf{p} \geq 2$ . The discrete space  $V_h^0$  must fulfill another condition, which is related to the constraint  $\nabla u \cdot \mathbf{n} = 0$  on  $\partial\Omega$ . Now, since the first derivative of a B-Spline curve is only defined by the two first control points, having  $u = 0$  and  $\nabla u \cdot \mathbf{n} = 0$  on the boundary is equivalent to set the first and last two control points to be exactly 0. Hence, our discrete space is defined as

$$V_h^0 = \text{span}\{N_i^p; 2 \leq i \leq \mathbf{n}-2\} \quad (14.23)$$

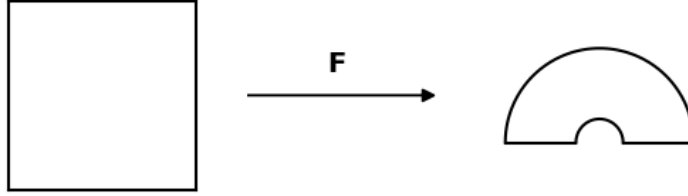


Figure 14.3: Coarse mesh in the logical and physical domain induced by a quadratic representation using two Bézier patches for each quarter of the annulus. In the radial direction, linear splines are used.

The discrete variational formulation is: Find  $u \in V_h^0$  such that

$$\int_{\hat{\Omega}} \Delta u \cdot \Delta v \, d\boldsymbol{\eta} = \int_{\hat{\Omega}} f v \, d\boldsymbol{\eta}, \quad \forall v \in V_h^0 \quad (14.24)$$

Let  $u_h \in V_h^0$  such that  $u_h = \sum_j u_j N_{\boldsymbol{\mu},j}^p$ . Then the weak formulation for the Biharmonic equation writes

$$\sum_j u_j \int_{\hat{\Omega}} \Delta N_{\boldsymbol{\mu},j}^p \cdot \Delta N_{\boldsymbol{\mu},i}^p \, d\boldsymbol{\eta} = \int_{\hat{\Omega}} f N_{\boldsymbol{\mu},i}^p \, d\boldsymbol{\eta}, \quad \forall 1 \leq i \leq n-1$$

In practice, in order to store these terms, we will need to have a reordering of the indices so that we map the set of multi-indices  $\{2 \leq \mathbf{i} \leq \mathbf{n}-2\}$  onto a set of 1d indices  $\{1 \leq I \leq N\}$ . Using this reordering, we get the following linear system

$$MU = F$$

where  $U$  denotes the coefficients ( $u_J$ ,  $1 \leq J \leq N$ ) and  $F$  is a vector consisting of the terms  $\int_{\hat{\Omega}} f N_{\boldsymbol{\mu},i}^p \, d\boldsymbol{\eta}$  for  $2 \leq i \leq n-2$ . Finally, the matrix  $M$  entries are defined as

$$M_{IJ} = \int_{\hat{\Omega}} \Delta N_{\boldsymbol{\mu},j}^p \cdot \Delta N_{\boldsymbol{\mu},i}^p \, d\boldsymbol{\eta}$$

### 14.2.2 Case with mapping

In this case, the coarse mesh is given by the mapping  $\mathbf{F}$  if it is discrete (constructed using B-Splines or NURBS). Let us denote  $V_{\mathbf{F}}$  the Schoenberg space associated to the mapping  $\mathbf{F}$ . We shall denote  $\{T_{\mathbf{F}}^{(\alpha)}, 1 \leq \alpha \leq d\}$  the knot vectors associated  $V_{\mathbf{F}}$ , the corresponding multiplicities to the beakpoints will be denoted  $\boldsymbol{\mu}_{\mathbf{F}}^{\alpha}$ , for every  $1 \leq \alpha \leq d$ .

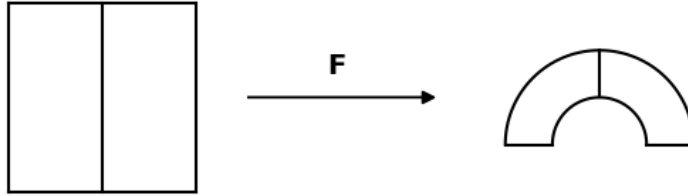


Figure 14.4: Coarse mesh in the logical and physical domain induced by a cubic representation using one Bézier patch. In the radial direction, linear splines are used.

The family of discrete subspaces needed for the Galerkin approximation, must be created such that the breakpoints of  $T_{\mathbf{F}}^{(\alpha)}$  are contained in  $T_N^{(\alpha)}$ , for every  $1 \leq \alpha \leq d$ . There are two reasons for this constraint:

1. A theoretical one, due to the definition of the bent Sobolev spaces,
2. A computational one, due to the quadrature rule that will be used on each element  $Q \in \mathcal{Q}_h(\hat{\Omega})$  of the logical domain.

In figure 14.3, we show a parametrization of half of the annulus using quadratic B-Splines (NURBS) in the angular direction, with the knot vector  $\{0, 0, 0, \frac{1}{2}, \frac{1}{2}, 1, 1, 1\}$ . In the radial direction, we use linear B-Splines with the knot vector  $\{0, 0, 1, 1\}$ . This leads to a coarse mesh of two elements in the first parametric direction and one in the second direction.

In figure 14.4, we show a parametrization of half of the annulus using cubic B-Splines (NURBS) in the angular direction, with the knot vector  $\{0, 0, 0, 0, 1, 1, 1, 1\}$ . In the radial direction, we use linear B-Splines with the knot vector  $\{0, 0, 1, 1\}$ . This leads to a coarse mesh of one element in each parametric direction.

Given a coarse parametrization of the domain, we can use the *h-refinement* strategy which is achieved by inserting new knots in each direction.

### Matrix form for Poisson

The discrete subspaces are defined as

$$V_h^0 = \text{span}\{N_i^p \circ \mathbf{F}^{-1}; 1 \leq i \leq n-1\} \quad (14.25)$$

where  $\mathbf{n} = (n_1, \dots, n_d)$ .

The discrete variational formulation is: Find  $u \in V_h^0$  such that

$$\int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x}, \quad \forall v \in V_h^0 \quad (14.26)$$

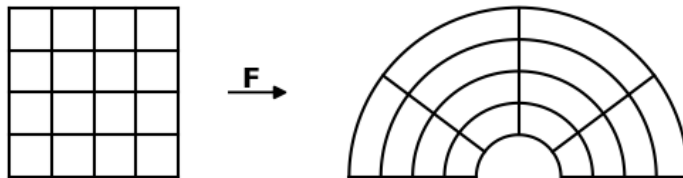


Figure 14.5: The mapping used in Fig. 14.4 after inserting 3 new knots in each direction, leading to 4 elements in each direction.

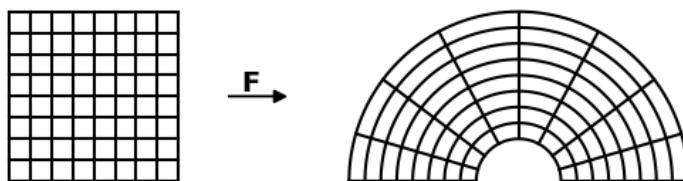


Figure 14.6: The mapping used in Fig. 14.4 after inserting 7 new knots in each direction, leading to 8 elements in each direction.

The next step is to map back the gradients from the physical domain to the logical one. Since for any function  $u$  we have,

$$\nabla u(\mathbf{x}) = (\mathcal{D}\mathbf{F})^{-T} \hat{\nabla} u(\mathbf{F}^{-1}(\mathbf{x})) \quad (14.27)$$

we can plug this expression in (14.26) which will lead to a kind of general elliptic problem on the logical domain. It is important to notice that for performance issues, we will not invert the jacobian matrix everytime, but rather computing the inverse analytically (or symbolically) in 2d and 3d.

**Exercise 14.2** Find the matrix form for the Biharmonic problem in the case where a mapping is used. ■

### 14.3 Exact DeRham sequences

In the sequel, we shall give exact DeRham sequences based on B-Splines in  $\mathbb{R}^d$  for  $d = 1, 2, 3$ . We assume that the domain  $\Omega$  is **simply connected**. Let us first recall that in 1D, DeRham sequence is reduced to

$$H^1(\Omega) \xrightarrow{\nabla} L^2(\Omega)$$

in 2d, we have two sequences

$$H^1(\Omega) \xrightarrow{\nabla} \mathbf{H}(\text{curl}, \Omega) \xrightarrow{\nabla \times} L^2(\Omega)$$

and

$$H^1(\Omega) \xrightarrow{\nabla \times} \mathbf{H}(\text{div}, \Omega) \xrightarrow{\nabla \cdot} L^2(\Omega)$$

while in 3d, we have the following sequence

$$H^1(\Omega) \xrightarrow{\nabla} \mathbf{H}(\text{curl}, \Omega) \xrightarrow{\nabla \times} \mathbf{H}(\text{div}, \Omega) \xrightarrow{\nabla \cdot} L^2(\Omega)$$

For simplicity, we shall introduce the operator  $d$ , that corresponds to the appropriate operator ( $\nabla$ ,  $\nabla \times$ , ...) for a given space, and rewrite our diagrams as

- 1D case:

$$H^1(\Omega) \xrightarrow{d} L^2(\Omega) \quad (14.28)$$

- 2D case:

$$H^1(\Omega) \xrightarrow{d} \mathbf{H}(\text{curl}, \Omega) \xrightarrow{d} L^2(\Omega) \quad (14.29)$$

$$H^1(\Omega) \xrightarrow{d} \mathbf{H}(\text{div}, \Omega) \xrightarrow{d} L^2(\Omega) \quad (14.30)$$

- 3D case:

$$H^1(\Omega) \xrightarrow{d} \mathbf{H}(\text{curl}, \Omega) \xrightarrow{d} \mathbf{H}(\text{div}, \Omega) \xrightarrow{d} L^2(\Omega) \quad (14.31)$$

### Exact sequence in 1D

We start by taking a subspace  $V_h(\text{grad}, \Omega) \subset H^1(\Omega)$ . In 1D,  $V_h(\text{grad}, \Omega)$  is a Schoenberg space associated to a given knot sequence  $T$ . We know that both B-Splines and M-Splines can be used as basis for the Schoenberg space. If we take the B-Splines as the canonical basis on  $V_h(\text{grad}, \Omega) := \text{span}\{N_i^p, 0 \leq i \leq n\}$ , we know that every function  $u \in V_h(\text{grad}, \Omega)$  can be written as  $u = \sum_i u_i N_i^p$ . By taking its derivative, we get

$$u' = \sum_i u_i (N_i^p)' = \sum_i (-u_{i-1} + u_i) M_i^{p-1}$$

since

$$(N_i^p)' = M_i^{p-1} - M_{i+1}^{p-1}$$

Now if we introduce the space  $V_h(L^2, \Omega) := \text{span}\{M_i^{p-1}, 0 \leq i \leq n\}$ . We just have shown that

$$\forall u \in V_h(\text{grad}, \Omega), \quad u' \in V_h(L^2, \Omega) \quad (14.32)$$

On the other hand, it is clear that the operator  $d$  is surjective (for any function  $u \in V_h(L^2, \Omega)$ , the function  $\int_{x_0}^x u(t) dt$  is clearly in  $V_h(\text{grad}, \Omega)$ ). This is summarized in the following diagram

$$V_h(\text{grad}, \Omega) \xrightarrow{d} V_h(L^2, \Omega) \quad (14.33)$$

where

$$\begin{cases} V_h(\text{grad}, \Omega) &= \mathcal{S}_{\boldsymbol{\mu}}^p &:= \text{span}\{N_i^p, 0 \leq i \leq n\} \\ V_h(L^2, \Omega) &= \mathcal{M}_{\boldsymbol{\mu}-1}^{p-1} &:= \text{span}\{M_i^{p-1}, 0 \leq i \leq n\} \end{cases} \quad (14.34)$$

**R** Let us introduce the B-Splines coefficients vector  $\mathbf{u} := (u_i)_{1 \leq i \leq n}$  (and  $\mathbf{u}^*$  for the derivative), we have

$$\mathbf{u}^* = \mathcal{D}\mathbf{u} \quad (14.35)$$

where  $\mathcal{D}$  is the incidence matrix (of entries  $-1$  and  $+1$ ).

### Exact sequences in 2D

Let us introduce the spaces

$$\begin{cases} V_h(\text{grad}, \Omega) &= \mathcal{S}_{\boldsymbol{\mu}_1}^{p_1} \otimes \mathcal{S}_{\boldsymbol{\mu}_2}^{p_2} \\ \mathbf{V}_h(\text{curl}, \Omega) &= \begin{pmatrix} \mathcal{M}_{\boldsymbol{\mu}_1-1}^{p_1-1} \otimes \mathcal{S}_{\boldsymbol{\mu}_2}^{p_2} \\ \mathcal{S}_{\boldsymbol{\mu}_1}^{p_1} \otimes \mathcal{M}_{\boldsymbol{\mu}_2-1}^{p_2-1} \end{pmatrix} \\ \mathbf{V}_h(\text{div}, \Omega) &= \begin{pmatrix} \mathcal{S}_{\boldsymbol{\mu}_1}^{p_1} \otimes \mathcal{M}_{\boldsymbol{\mu}_2-1}^{p_2-1} \\ \mathcal{M}_{\boldsymbol{\mu}_1-1}^{p_1-1} \otimes \mathcal{S}_{\boldsymbol{\mu}_2}^{p_2} \end{pmatrix} \\ V_h(L^2, \Omega) &= \mathcal{M}_{\boldsymbol{\mu}_1-1}^{p_1-1} \otimes \mathcal{M}_{\boldsymbol{\mu}_2-1}^{p_2-1} \end{cases} \quad (14.36)$$

**Lemma 14** We have the following properties

- $\nabla V_h(\text{grad}, \Omega) \subset \mathbf{V}_h(\text{curl}, \Omega)$
- $\nabla \times \mathbf{V}_h(\text{curl}, \Omega) \subset V_h(L^2, \Omega)$



- $\nabla \times V_h(\text{grad}, \Omega) \subset \mathbf{V}_h(\text{curl}, \Omega)$
- $\nabla \cdot \mathbf{V}_h(\text{curl}, \Omega) \subset V_h(L^2, \Omega)$

*Proof.* Can be deduced thanks to the property (14.32). ■

**Proposition 14.3.1** We have the following diagrams

$$V_h(\text{grad}, \Omega) \xrightarrow{d} \mathbf{V}_h(\text{curl}, \Omega) \xrightarrow{d} V_h(L^2, \Omega) \quad (14.37)$$

$$V_h(\text{grad}, \Omega) \xrightarrow{d} \mathbf{V}_h(\text{div}, \Omega) \xrightarrow{d} V_h(L^2, \Omega) \quad (14.38)$$

*Proof.* We shall restrict our proof to the sequence (14.37) and show that using the previous discrete spaces, we maintain the same relationships as in the continuous case.

Since  $\Omega$  is simply connected, we know that in the continuous case, the kernel of the curl operator is equal to the range of the gradient operator, *i.e.*

$$\nabla(H^1(\Omega)) = \mathcal{N}(\nabla \times)$$

Showing the exactness of the discrete DeRham sequence means that the last property is preserved in the discrete case. We shall then show that

$$\nabla(V_h(\text{grad}, \Omega)) = \mathcal{N}(\nabla \times) \cap \mathbf{V}_h(\text{curl}, \Omega)$$

since for any  $u \in V_h(\text{grad}, \Omega)$  we have  $\nabla \times \nabla u = 0$ , we get

$$\nabla(V_h(\text{grad}, \Omega)) \subset \mathcal{N}(\nabla \times) \cap \mathbf{V}_h(\text{curl}, \Omega)$$

In order to show the equality, it is enough to show that both spaces have the same dimension. But we have  $\dim V_h(\text{grad}, \Omega) = (n_1 + 1)(n_2 + 1)$ . But since,  $\Omega$  is simply connected we have  $\dim \mathcal{N}(\nabla) = 1$  and therefore we have

$$\dim \nabla(V_h(\text{grad}, \Omega)) = (n_1 + 1)(n_2 + 1) - 1$$

On the other hand, we have

$$\begin{aligned} \dim \mathbf{V}_h(\text{curl}, \Omega) &= \dim(\mathcal{N}(\nabla \times) \cap \mathbf{V}_h(\text{curl}, \Omega)) + \dim(\mathcal{R}(\nabla \times) \cap \mathbf{V}_h(\text{curl}, \Omega)) \\ &= n_1(n_2 + 1) + (n_1 + 1)n_2 \end{aligned}$$

meaning

$$\dim(\mathcal{N}(\nabla \times) \cap \mathbf{V}_h(\text{curl}, \Omega)) = 2n_1n_2 + n_1 + n_2 - \dim(\mathcal{R}(\nabla \times) \cap \mathbf{V}_h(\text{curl}, \Omega))$$

In order to show that

$$\dim \nabla(V_h(\text{grad}, \Omega)) = \dim(\mathcal{N}(\nabla \times) \cap \mathbf{V}_h(\text{curl}, \Omega))$$

it is then enough to show that

$$\dim(\mathcal{R}(\nabla \times) \cap \mathbf{V}_h(\text{curl}, \Omega)) = n_1n_2$$

but since the operator  $\nabla \times : \mathbf{V}_h(\text{curl}, \Omega) \rightarrow V_h(L^2, \Omega)$  is surjective (proof similar to the 1d case), we deduce that  $\nabla \times (\mathbf{V}_h(\text{curl}, \Omega)) = V_h(L^2, \Omega)$ . Finally, we have  $\dim V_h(L^2, \Omega) = n_1n_2$  which completes the proof. ■

**R** The discrete derivatives in 2D are given by

$$\begin{cases} \mathbf{G} = \begin{pmatrix} \mathcal{D} \otimes \mathbb{I} \\ \mathbb{I} \otimes \mathcal{D} \end{pmatrix} \\ \mathbf{C} = \begin{pmatrix} \mathbb{I} \otimes \mathcal{D} \\ -\mathcal{D} \otimes \mathbb{I} \end{pmatrix} \\ \mathbf{C} = \begin{pmatrix} -\mathbb{I} \otimes \mathcal{D} & \mathcal{D} \otimes \mathbb{I} \end{pmatrix} \\ \mathbf{D} = \begin{pmatrix} \mathcal{D} \otimes \mathbb{I} & \mathbb{I} \otimes \mathcal{D} \end{pmatrix} \end{cases} \quad (14.39)$$

where we used the identity matrix  $\mathbb{I}$ .

### Exact sequence in 3D

Let us introduce the spaces

$$\begin{cases} V_h(\text{grad}, \Omega) = \mathcal{S}_{\boldsymbol{\mu}_1}^{p_1} \otimes \mathcal{S}_{\boldsymbol{\mu}_2}^{p_2} \otimes \mathcal{S}_{\boldsymbol{\mu}_3}^{p_3} \\ \mathbf{V}_h(\text{curl}, \Omega) = \begin{pmatrix} \mathcal{M}_{\boldsymbol{\mu}_1-1}^{p_1-1} \otimes \mathcal{S}_{\boldsymbol{\mu}_2}^{p_2} \otimes \mathcal{S}_{\boldsymbol{\mu}_3}^{p_3} \\ \mathcal{S}_{\boldsymbol{\mu}_1}^{p_1} \otimes \mathcal{M}_{\boldsymbol{\mu}_2-1}^{p_2-1} \otimes \mathcal{S}_{\boldsymbol{\mu}_3}^{p_3} \\ \mathcal{S}_{\boldsymbol{\mu}_1}^{p_1} \otimes \mathcal{S}_{\boldsymbol{\mu}_2}^{p_2} \otimes \mathcal{M}_{\boldsymbol{\mu}_3-1}^{p_3-1} \end{pmatrix} \\ \mathbf{V}_h(\text{div}, \Omega) = \begin{pmatrix} \mathcal{S}_{\boldsymbol{\mu}_1}^{p_1} \otimes \mathcal{M}_{\boldsymbol{\mu}_2-1}^{p_2-1} \otimes \mathcal{M}_{\boldsymbol{\mu}_3-1}^{p_3-1} \\ \mathcal{M}_{\boldsymbol{\mu}_1-1}^{p_1-1} \otimes \mathcal{S}_{\boldsymbol{\mu}_2}^{p_2} \otimes \mathcal{M}_{\boldsymbol{\mu}_3-1}^{p_3-1} \\ \mathcal{M}_{\boldsymbol{\mu}_1-1}^{p_1-1} \otimes \mathcal{M}_{\boldsymbol{\mu}_2-1}^{p_2-1} \otimes \mathcal{S}_{\boldsymbol{\mu}_3}^{p_3} \end{pmatrix} \\ V_h(L^2, \Omega) = \mathcal{M}_{\boldsymbol{\mu}_1-1}^{p_1-1} \otimes \mathcal{M}_{\boldsymbol{\mu}_2-1}^{p_2-1} \otimes \mathcal{M}_{\boldsymbol{\mu}_3-1}^{p_3-1} \end{cases} \quad (14.40)$$

**Proposition 14.3.2** We have the following diagram

$$V_h(\text{grad}, \Omega) \xrightarrow{\mathbf{d}} \mathbf{V}_h(\text{curl}, \Omega) \xrightarrow{\mathbf{d}} \mathbf{V}_h(\text{div}, \Omega) \xrightarrow{\mathbf{d}} V_h(L^2, \Omega) \quad (14.41)$$

**R** The discrete derivatives in 3D are given by

$$\begin{cases} \mathbf{G} = \begin{pmatrix} \mathcal{D} \otimes \mathbb{I} \otimes \mathbb{I} \\ \mathbb{I} \otimes \mathcal{D} \otimes \mathbb{I} \\ \mathbb{I} \otimes \mathbb{I} \otimes \mathcal{D} \end{pmatrix} \\ \mathbf{C} = \begin{pmatrix} \mathbf{0} & -\mathbb{I} \otimes \mathbb{I} \otimes \mathcal{D} & \mathbb{I} \otimes \mathcal{D} \otimes \mathbb{I} \\ \mathbb{I} \otimes \mathbb{I} \otimes \mathcal{D} & \mathbf{0} & -\mathcal{D} \otimes \mathbb{I} \otimes \mathbb{I} \\ -\mathbb{I} \otimes \mathcal{D} \otimes \mathbb{I} & \mathcal{D} \otimes \mathbb{I} \otimes \mathbb{I} & \mathbf{0} \end{pmatrix} \\ \mathbf{D} = \begin{pmatrix} \mathcal{D} \otimes \mathbb{I} \otimes \mathbb{I} & \mathbb{I} \otimes \mathcal{D} \otimes \mathbb{I} & \mathbb{I} \otimes \mathbb{I} \otimes \mathcal{D} \end{pmatrix} \end{cases} \quad (14.42)$$

where we used the identity matrix  $\mathbb{I}$ .

### 14.3.1 Commuting diagrams

DeRham diagrams such as (14.28), (14.29), (14.30) and (14.31) can be extended with additional operators. Among the important properties, one can build specific projectors that make these diagrams commute. We shall start with the 1D case, then extend it by tensor product. For this purpose, we take any interpolator, denoted by  $I_T^p$ , (Fix-DeBoor for theoretical studies, or B-Splines interpolator), quasi-interpolator or least-square approximation associated to a given knot sequence  $T$  and degree  $p$ , and then we define the corresponding histopolation operator by

$$H_T^p f := DI_T^{p+1} \left( x \mapsto \int_{x_0}^x f dx \right) \quad (14.43)$$

**1D case**

We then introduce the operators

$$\begin{cases} P_h^{\text{grad}} & := I_T^p \\ P_h^{L^2} & := H_T^{p-1} \end{cases} \quad (14.44)$$

**Proposition 14.3.3** For any  $u \in C^\infty(\Omega)$ , we have

$$dP_h^{\text{grad}}(u) = P_h^{L^2}(du) \quad (14.45)$$

**Exercise 14.3** Prove the previous proposition. ■

These results are summarized in the following commuting diagram

$$\begin{array}{ccc} H^s(\Omega) & \xrightarrow{d} & H^{s-1}(\Omega) \\ P_h^{\text{grad}} \downarrow & & P_h^{L^2} \downarrow \\ V_h(\text{grad}, \Omega) & \xrightarrow{\mathbb{G}} & V_h(L^2, \Omega) \end{array} \quad (14.46)$$

**2D case**

We then introduce the operators

$$\begin{cases} P_h^{\text{grad}} & := I_{T_1}^{p_1} \odot I_{T_2}^{p_2} \\ P_h^{\text{curl}} & := \begin{pmatrix} H_{T_1}^{p_1-1} \odot I_{T_2}^{p_2} \\ I_{T_1}^{p_1} \odot H_{T_2}^{p_2-1} \end{pmatrix} \\ P_h^{\text{div}} & := \begin{pmatrix} I_{T_1}^{p_1} \odot H_{T_2}^{p_2-1} \\ H_{T_1}^{p_1-1} \odot I_{T_2}^{p_2} \end{pmatrix} \\ P_h^{L^2} & := H_{T_1}^{p_1-1} \odot H_{T_2}^{p_2-1} \end{cases} \quad (14.47)$$

where the notation  $\odot$  to express the composition of operators on each coordinate, for example,

$$I_{T_1}^{p_1} \odot I_{T_2}^{p_2} f := I_{T_1}^{p_1}(x \mapsto I_{T_2}^{p_2}(y \mapsto f(x, y)))$$

**Proposition 14.3.4** For any  $u \in C^\infty(\Omega)$  and  $\mathbf{u} \in \mathbf{C}^\infty(\Omega)$ , we have

1) for the first sequence (14.37)

$$dP_h^{\text{grad}}(\mathbf{u}) = P_h^{\text{curl}}(d\mathbf{u}) \quad (14.48)$$

$$dP_h^{\text{curl}}(\mathbf{u}) = P_h^{L^2}(d\mathbf{u}) \quad (14.49)$$

2) for the second sequence (14.38)

$$dP_h^{\text{grad}}(\mathbf{u}) = P_h^{\text{div}}(d\mathbf{u}) \quad (14.50)$$

$$dP_h^{\text{div}}(\mathbf{u}) = P_h^{L^2}(d\mathbf{u}) \quad (14.51)$$

**Exercise 14.4** Prove the previous proposition. ■

These results are summarized in the following commuting diagrams

$$\begin{array}{ccccc}
 H^1(\Omega) & \xrightarrow{\nabla} & \mathbf{H}(\text{curl}, \Omega) & \xrightarrow{\nabla \times} & L^2(\Omega) \\
 P_h^{\text{grad}} \downarrow & & P_h^{\text{curl}} \downarrow & & P_h^{L^2} \downarrow \\
 V_h(\text{grad}, \Omega) & \xrightarrow{\mathbb{G}} & \mathbf{V}_h(\text{curl}, \Omega) & \xrightarrow{\mathbb{C}} & V_h(L^2, \Omega)
 \end{array} \tag{14.52}$$

and

$$\begin{array}{ccccc}
 H^1(\Omega) & \xrightarrow{\nabla \times} & \mathbf{H}(\text{div}, \Omega) & \xrightarrow{\nabla \cdot} & L^2(\Omega) \\
 P_h^{\text{grad}} \downarrow & & P_h^{\text{div}} \downarrow & & P_h^{L^2} \downarrow \\
 V_h(\text{grad}, \Omega) & \xrightarrow{\mathbb{C}} & \mathbf{V}_h(\text{div}, \Omega) & \xrightarrow{\mathbb{D}} & V_h(L^2, \Omega)
 \end{array} \tag{14.53}$$

### 3D case

We then introduce the operators

$$\left\{ \begin{array}{l}
 P_h^{\text{grad}} := I_{T_1}^{p_1} \circ I_{T_2}^{p_2} \circ I_{T_3}^{p_3} \\
 P_h^{\text{curl}} := \left( \begin{array}{l}
 H_{T_1}^{p_1-1} \circ I_{T_2}^{p_2} \circ I_{T_3}^{p_3} \\
 I_{T_1}^{p_1} \circ H_{T_2}^{p_2-1} \circ I_{T_3}^{p_3} \\
 I_{T_1}^{p_1} \circ I_{T_2}^{p_2} \circ H_{T_3}^{p_3-1}
 \end{array} \right) \\
 P_h^{\text{div}} := \left( \begin{array}{l}
 I_{T_1}^{p_1} \circ I_{T_2}^{p_2} \circ H_{T_3}^{p_3-1} \\
 I_{T_1}^{p_1} \circ H_{T_2}^{p_2-1} \circ I_{T_3}^{p_3} \\
 H_{T_1}^{p_1-1} \circ I_{T_2}^{p_2} \circ I_{T_3}^{p_3}
 \end{array} \right) \\
 P_h^{L^2} := H_{T_1}^{p_1-1} \circ H_{T_2}^{p_2-1} \circ H_{T_3}^{p_3-1}
 \end{array} \right. \tag{14.54}$$

**Proposition 14.3.5** For any  $u \in C^\infty(\Omega)$  and  $\mathbf{u} \in \mathbf{C}^\infty(\Omega)$ , we have

$$d P_h^{\text{grad}}(u) = P_h^{\text{curl}}(d u) \tag{14.55}$$

$$d P_h^{\text{curl}}(\mathbf{u}) = P_h^{\text{div}}(d \mathbf{u}) \tag{14.56}$$

$$d P_h^{\text{div}}(\mathbf{u}) = P_h^{L^2}(d \mathbf{u}) \tag{14.57}$$

**Exercise 14.5** Prove the previous proposition. ■

These results are summarized in the following commuting diagram

$$\begin{array}{ccccccc}
 H^1(\Omega) & \xrightarrow{\nabla} & \mathbf{H}(\text{curl}, \Omega) & \xrightarrow{\nabla \times} & \mathbf{H}(\text{div}, \Omega) & \xrightarrow{\nabla \cdot} & L^2(\Omega) \\
 P_h^{\text{grad}} \downarrow & & P_h^{\text{curl}} \downarrow & & P_h^{\text{div}} \downarrow & & P_h^{L^2} \downarrow \\
 V_h(\text{grad}, \Omega) & \xrightarrow{\mathbb{G}} & \mathbf{V}_h(\text{curl}, \Omega) & \xrightarrow{\mathbb{C}} & \mathbf{V}_h(\text{div}, \Omega) & \xrightarrow{\mathbb{D}} & V_h(L^2, \Omega)
 \end{array} \tag{14.58}$$

### 14.3.2 Examples without mapping

#### Matrix form for $H(\text{curl}, \Omega)$ -elliptic problem:

We consider the discrete variational formulation of (12.18), for which the associated matrix form was given in (12.2.4). Here  $V_h := \mathbf{V}_h(\text{curl}, \Omega)$ . As in opposition to the previous chapter, we now have an explicit expression for the basis functions, we shall use them to get a final matrix form that can be implemented. For the sake of simplicity we shall introduce the scalar functions

$$\begin{aligned}\Psi_j^1 &= M_{\mu_1-1, i_1}^{p_1-1} N_{\mu_2, i_2}^{p_2} N_{\mu_3, i_3}^{p_3} \\ \Psi_j^2 &= N_{\mu_1, i_1}^{p_1} M_{\mu_2-1, i_2}^{p_2-1} N_{\mu_3, i_3}^{p_3} \\ \Psi_j^3 &= N_{\mu_1, i_1}^{p_1} N_{\mu_2, i_2}^{p_2} M_{\mu_3-1, i_3}^{p_3-1}\end{aligned}$$

we also define the vectors  $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ,  $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  and  $\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ . Therefore, the expression of  $\mathbf{u}_h \in \mathbf{V}_h(\text{curl}, \Omega)$  becomes

$$\mathbf{u}_h = \sum_j (u_j^1 \Psi_j^1 \mathbf{e}_1 + u_j^2 \Psi_j^2 \mathbf{e}_2 + u_j^3 \Psi_j^3 \mathbf{e}_3)$$

On the other hand, we have,

$$\begin{aligned}\nabla \times \Psi_j^1 \mathbf{e}_1 &= \partial_{x_3} \Psi_j^1 \mathbf{e}_2 - \partial_{x_2} \Psi_j^1 \mathbf{e}_3 \\ \nabla \times \Psi_j^2 \mathbf{e}_2 &= -\partial_{x_3} \Psi_j^2 \mathbf{e}_1 + \partial_{x_1} \Psi_j^2 \mathbf{e}_3 \\ \nabla \times \Psi_j^3 \mathbf{e}_3 &= \partial_{x_2} \Psi_j^3 \mathbf{e}_1 - \partial_{x_1} \Psi_j^3 \mathbf{e}_2\end{aligned}$$

Because  $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ , we get,

$$\begin{aligned}\nabla \times \Psi_j^1 \mathbf{e}_1 \cdot \nabla \times \Psi_i^1 \mathbf{e}_1 &= \partial_{x_3} \Psi_j^1 \partial_{x_3} \Psi_i^1 + \partial_{x_2} \Psi_j^1 \partial_{x_2} \Psi_i^1 \\ \nabla \times \Psi_j^1 \mathbf{e}_1 \cdot \nabla \times \Psi_i^2 \mathbf{e}_2 &= -\partial_{x_2} \Psi_j^1 \partial_{x_1} \Psi_i^2 \\ \nabla \times \Psi_j^1 \mathbf{e}_1 \cdot \nabla \times \Psi_i^3 \mathbf{e}_3 &= -\partial_{x_3} \Psi_j^1 \partial_{x_1} \Psi_i^3 \\ \nabla \times \Psi_j^2 \mathbf{e}_2 \cdot \nabla \times \Psi_i^2 \mathbf{e}_2 &= \partial_{x_3} \Psi_j^2 \partial_{x_3} \Psi_i^2 + \partial_{x_1} \Psi_j^2 \partial_{x_1} \Psi_i^2 \\ \nabla \times \Psi_j^2 \mathbf{e}_2 \cdot \nabla \times \Psi_i^3 \mathbf{e}_3 &= -\partial_{x_3} \Psi_j^2 \partial_{x_2} \Psi_i^3 \\ \nabla \times \Psi_j^3 \mathbf{e}_3 \cdot \nabla \times \Psi_i^3 \mathbf{e}_3 &= \partial_{x_2} \Psi_j^3 \partial_{x_2} \Psi_i^3 + \partial_{x_1} \Psi_j^3 \partial_{x_1} \Psi_i^3\end{aligned}$$

we find that  $A$  is a symmetric  $3 \times 3$  block matrix of the form

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{12}^T & A_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33} \end{pmatrix} \quad (14.59)$$

where

$$\begin{aligned}
A_{11i,j} &= \int_{\Omega} \partial_{x_3} \Psi_j^1 \partial_{x_3} \Psi_i^1 + \partial_{x_2} \Psi_j^1 \partial_{x_2} \Psi_i^1 \, dx + \int_{\Omega} \mu \Psi_j^1 \Psi_i^1 \, dx \\
A_{12i,j} &= - \int_{\Omega} \partial_{x_2} \Psi_j^1 \partial_{x_1} \Psi_i^2 \, dx \\
A_{13i,j} &= - \int_{\Omega} \partial_{x_3} \Psi_j^1 \partial_{x_1} \Psi_i^3 \, dx \\
A_{22i,j} &= \int_{\Omega} \partial_{x_3} \Psi_j^2 \partial_{x_3} \Psi_i^2 + \partial_{x_1} \Psi_j^2 \partial_{x_1} \Psi_i^2 \, dx + \int_{\Omega} \mu \Psi_j^2 \Psi_i^2 \, dx \\
A_{23i,j} &= - \int_{\Omega} \partial_{x_3} \Psi_j^2 \partial_{x_2} \Psi_i^3 \, dx \\
A_{33i,j} &= \int_{\Omega} \partial_{x_2} \Psi_j^3 \partial_{x_2} \Psi_i^3 + \partial_{x_1} \Psi_j^3 \partial_{x_1} \Psi_i^3 \, dx + \int_{\Omega} \mu \Psi_j^3 \Psi_i^3 \, dx
\end{aligned}$$

For the right hand side, the entries associated to each component of the vector  $\mathbf{f}$  are given by

$$\begin{aligned}
F_{1,i} &= \int_{\Omega} \mathbf{f}_1 \Psi_i^1 \, dx \\
F_{2,i} &= \int_{\Omega} \mathbf{f}_2 \Psi_i^2 \, dx \\
F_{3,i} &= \int_{\Omega} \mathbf{f}_3 \Psi_i^3 \, dx
\end{aligned}$$

**Matrix form for  $H(\operatorname{div}, \Omega)$ -elliptic problem:**

**Exercise 14.6** Compute the matrix form for the  $H(\operatorname{div}, \Omega)$ -elliptic problem. ■

**Matrix form for Poisson using mixed FEM:**

We consider the discrete variational formulation of (12.38), for which the associated matrix form was given in (12.3.3). Here  $V_h := \mathbf{V}_h(\operatorname{div}, \Omega)$  and  $W_h := V_h(L^2, \Omega)$ . As in opposition to the previous chapter, we now have an explicit expression for the basis functions, we shall use them to get a final matrix form that can be implemented. For the sake of simplicity we shall introduce the scalar functions

$$\begin{aligned}
\Psi_j^1 &= N_{\boldsymbol{\mu}_{1,i_1}}^{p_1} M_{\boldsymbol{\mu}_{2-1,i_2}}^{p_2-1} M_{\boldsymbol{\mu}_{3-1,i_3}}^{p_3-1} \\
\Psi_j^2 &= M_{\boldsymbol{\mu}_{1-1,i_1}}^{p_1-1} N_{\boldsymbol{\mu}_{2,i_2}}^{p_2} M_{\boldsymbol{\mu}_{3-1,i_3}}^{p_3-1} \\
\Psi_j^3 &= M_{\boldsymbol{\mu}_{1-1,i_1}}^{p_1-1} M_{\boldsymbol{\mu}_{2-1,i_2}}^{p_2-1} N_{\boldsymbol{\mu}_{3,i_3}}^{p_3}
\end{aligned}$$

and

$$\phi_j = M_{\boldsymbol{\mu}_{1-1,i_1}}^{p_1-1} M_{\boldsymbol{\mu}_{2-1,i_2}}^{p_2-1} M_{\boldsymbol{\mu}_{3-1,i_3}}^{p_3-1}$$

we also define the vectors  $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ,  $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  and  $\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ . Therefore, the expression of

$\mathbf{u}_h \in \mathbf{V}_h(\operatorname{div}, \Omega)$  becomes

$$\mathbf{u}_h = \sum_j (u_j^1 \Psi_j^1 \mathbf{e}_1 + u_j^2 \Psi_j^2 \mathbf{e}_2 + u_j^3 \Psi_j^3 \mathbf{e}_3)$$

Because  $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ , we find that  $A$  is a block diagonal matrix of the form

$$A = \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{pmatrix} \tag{14.60}$$

where

$$\begin{aligned} A_{11i,j} &= \int_{\Omega} \Psi_i^1 \Psi_j^1 \, d\mathbf{x} \\ A_{22i,j} &= \int_{\Omega} \Psi_i^2 \Psi_j^2 \, d\mathbf{x} \\ A_{33i,j} &= \int_{\Omega} \Psi_i^3 \Psi_j^3 \, d\mathbf{x} \end{aligned}$$

While for  $B$  we have

$$B = \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} \quad (14.61)$$

where

$$\begin{aligned} B_{1i,j} &= - \int_{\Omega} \phi_j \partial_{x_1} \Psi_i^1 \, d\mathbf{x} \\ B_{2i,j} &= - \int_{\Omega} \phi_j \partial_{x_2} \Psi_i^2 \, d\mathbf{x} \\ B_{3i,j} &= - \int_{\Omega} \phi_j \partial_{x_3} \Psi_i^3 \, d\mathbf{x} \end{aligned}$$

The final matrix form is

$$\begin{pmatrix} A_{11} & 0 & 0 & B_1 \\ 0 & A_{22} & 0 & B_2 \\ 0 & 0 & A_{33} & B_3 \\ B_1^T & B_2^T & B_3^T & 0 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ P \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ F \end{pmatrix}$$

with

$$F_i = - \int_{\Omega} f \phi_i \, d\mathbf{x}, \quad 1 \leq i \leq N_{W_h}$$

**Exercise 14.7** By exploiting the tensor structure of the basis, rewrite the matrix form in terms of Kronecker products. ■

**Exercise 14.8** Write the matrix form of the second formulation (12.3.3) of the Mixed Poisson problem. ■

#### Matrix form for the Stokes problem:

We consider the discrete variational formulation of (12.42), for which the associated matrix form was given in (12.3.3). Here  $V_h := V_h(\text{grad}, \Omega)$  and  $W_h := V_h(L^2, \Omega)$ . As in opposition to the previous chapter, we now have an explicit expression for the basis functions, we shall use them to get a final matrix form that can be implemented. For the sake of simplicity we shall introduce the scalar functions

$$\begin{aligned} \Psi_j^1 &= N_{\boldsymbol{\mu}_1, i_1}^{p_1} N_{\boldsymbol{\mu}_2, i_2}^{p_2} N_{\boldsymbol{\mu}_3, i_3}^{p_3} \\ \Psi_j^2 &= N_{\boldsymbol{\mu}_1, i_1}^{p_1} N_{\boldsymbol{\mu}_2, i_2}^{p_2} N_{\boldsymbol{\mu}_3, i_3}^{p_3} \\ \Psi_j^3 &= N_{\boldsymbol{\mu}_1, i_1}^{p_1} N_{\boldsymbol{\mu}_2, i_2}^{p_2} N_{\boldsymbol{\mu}_3, i_3}^{p_3} \end{aligned}$$

and

$$\phi_j = M_{\mu_1-1, i_1}^{p_1-1} M_{\mu_2-1, i_2}^{p_2-1} M_{\mu_3-1, i_3}^{p_3-1}$$

we also define the vectors  $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ,  $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  and  $\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ . Therefore, the expression of  $\mathbf{u}_h \in \mathbf{V}_h(\text{div}, \Omega)$  becomes

$$\mathbf{u}_h = \sum_j (u_j^1 \Psi_j^1 \mathbf{e}_1 + u_j^2 \Psi_j^2 \mathbf{e}_2 + u_j^3 \Psi_j^3 \mathbf{e}_3)$$

Because  $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ , we find that  $A$  is a block diagonal matrix of the form

$$A = \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{pmatrix} \quad (14.62)$$

where

$$A_{11i,j} = \int_{\Omega} \nabla \Psi_i^1 \cdot \nabla \Psi_j^1 \, dx$$

$$A_{22i,j} = \int_{\Omega} \nabla \Psi_i^2 \cdot \nabla \Psi_j^2 \, dx$$

$$A_{33i,j} = \int_{\Omega} \nabla \Psi_i^3 \cdot \nabla \Psi_j^3 \, dx$$

While for  $B$  we have

$$B = \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} \quad (14.63)$$

where

$$B_{1i,j} = - \int_{\Omega} \phi_j \partial_{x_1} \Psi_i^1 \, dx$$

$$B_{2i,j} = - \int_{\Omega} \phi_j \partial_{x_2} \Psi_i^2 \, dx$$

$$B_{3i,j} = - \int_{\Omega} \phi_j \partial_{x_3} \Psi_i^3 \, dx$$

The final matrix form is

$$\begin{pmatrix} A_{11} & 0 & 0 & B_1 \\ 0 & A_{22} & 0 & B_2 \\ 0 & 0 & A_{33} & B_3 \\ B_1^T & B_2^T & B_3^T & 0 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ P \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \\ 0 \end{pmatrix}$$

with

$$F_{1,i} = \int_{\Omega} \mathbf{f} \cdot \Psi_i^1 \, dx$$

$$F_{2,i} = \int_{\Omega} \mathbf{f} \cdot \Psi_i^2 \, dx$$

$$F_{3,i} = \int_{\Omega} \mathbf{f} \cdot \Psi_i^3 \, dx$$



**Exercise 14.9** By exploiting the tensor structure of the basis, rewrite the matrix form in terms of Kronecker products. ■

**Exercise 14.10** Write the matrix form of the second formulation (12.3.3) of the Mixed Poisson problem. ■

### 14.3.3 Pullbacks

In the case where the physical domain  $\Omega := \mathcal{F}(\hat{\Omega})$  is the *image* of a *logical* domain  $\hat{\Omega}$  by a smooth mapping  $\mathcal{F}$  (at least  $\mathcal{C}^1$ ), we have the following *parallel* diagrams (14.64)

$$\begin{array}{ccccccc}
 H^1(\Omega) & \xrightarrow{\nabla} & \mathbf{H}(\text{curl}, \Omega) & \xrightarrow{\nabla \times} & \mathbf{H}(\text{div}, \Omega) & \xrightarrow{\nabla \cdot} & L^2(\Omega) \\
 \uparrow \iota^0 & & \uparrow \iota^1 & & \uparrow \iota^2 & & \uparrow \iota^3 \\
 H^1(\hat{\Omega}) & \xrightarrow{\nabla} & H(\text{curl}, \hat{\Omega}) & \xrightarrow{\nabla \times} & H(\text{div}, \hat{\Omega}) & \xrightarrow{\nabla \cdot} & L^2(\hat{\Omega})
 \end{array} \quad (14.64)$$

Where the mappings  $\iota^0, \iota^1, \iota^2$  and  $\iota^3$  are called *pullbacks* and are given by Eq. (14.65), (14.66), (14.67) and (14.68). with  $D\mathcal{F}$  is the jacobian matrix of the mapping  $\mathcal{F}$ . We recall that the pullbacks

$$\iota^0 \hat{\phi} := \mathbf{x} \rightarrow \hat{\phi}(\mathbf{F}^{-1}(\mathbf{x})) \quad \forall \hat{\phi} \in H^1(\hat{\Omega}) \quad (14.65)$$

$$\iota^1 \hat{\Psi} := \mathbf{x} \rightarrow (\mathcal{D}\mathbf{F})^{-T} \hat{\Psi}(\mathbf{F}^{-1}(\mathbf{x})) \quad \forall \hat{\Psi} \in H(\text{curl}, \hat{\Omega}) \quad (14.66)$$

$$\iota^2 \hat{\Phi} := \mathbf{x} \rightarrow \frac{1}{J} \mathcal{D}\mathbf{F} \hat{\Phi}(\mathbf{F}^{-1}(\mathbf{x})) \quad \forall \hat{\Phi} \in H(\text{div}, \hat{\Omega}) \quad (14.67)$$

$$\iota^3 \hat{\rho} := \mathbf{x} \rightarrow \hat{\rho}(\mathbf{F}^{-1}(\mathbf{x})) \quad \forall \hat{\rho} \in L^2(\hat{\Omega}) \quad (14.68)$$

$\iota^0, \iota^1, \iota^2$  and  $\iota^3$  are *isomorphisms* between the corresponding spaces, with  $\iota^1$  and  $\iota^2$  are the Piola transformations.

**Exercise 14.11** Write these matrices in an explicit form, using symbolic calculus or analytically. ■

**14.3.4 Examples with mapping****Matrix form for  $H(\text{curl}, \Omega)$ -elliptic problem:**

TODO

**Matrix form for  $H(\text{div}, \Omega)$ -elliptic problem:**

TODO

**Matrix form for Poisson using mixed FEM:**

TODO

**14.4 Problems**

TODO

## **15. Historical Notes**

TODO



# Linear Algebra

<b>16</b>	<b>Kronecker Algebra .....</b>	<b>142</b>
16.1	Kronecker algebra	
16.2	Problems	
<b>17</b>	<b>Historical Notes .....</b>	<b>147</b>



## 16. Kronecker Algebra

### 16.1 Kronecker algebra

In this section, we present an overview about an interesting subject, which is the Kronecker Algebra, and which will be of a big interest in the Fast-IGA approach. Most of the presented results were taken from [11, 20].

**Definition 16.1.1 — The  $\text{vec}$  operator.** Let  $A = (a_{ij}) \in \mathcal{M}_{n \times m}$ , the  $\text{vec}$  operator is defined as,

$$\text{vec}A = \begin{pmatrix} A_{:,1} \\ \vdots \\ A_{:,m} \end{pmatrix} \in \mathbb{R}^{mn} \quad (16.1)$$

which is simply a vector composed by stacking all the columns of  $A$ . Where we denote  $A_{:,j}$  the  $j^{\text{th}}$  column of  $A$ .

We also define the inverse operator of  $\text{vec}$  by,

$$A = \text{vec}^{-1}\text{vec}A \quad (16.2)$$

**Definition 16.1.2 — Kronecker product.** Let  $A = (a_{ij}) \in \mathcal{M}_{m \times n}$  and  $B = (b_{ij}) \in \mathcal{M}_{r \times s}$  be two matrices. The Kronecker product of  $A$  and  $B$ , denoted by  $A \otimes B \in \mathcal{M}_{mr \times ns}$ , defines the following matrix:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix} \quad (16.3)$$

**Example**

Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

then their Kronecker product is,

$$A \otimes B = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix} \quad (16.4)$$

**Properties****Proposition 16.1.1** If  $\alpha$  is a scalar, then

$$A \otimes \alpha B = \alpha A \otimes B \quad (16.5)$$

**Proposition 16.1.2** We have,

$$(A + B) \otimes C = A \otimes C + B \otimes C \quad (16.6)$$

$$A \otimes (B + C) = A \otimes B + A \otimes C \quad (16.7)$$

**Proposition 16.1.3 — Associativity.**

$$A \otimes B \otimes C = A \otimes (B \otimes C) = (A \otimes B) \otimes C \quad (16.8)$$

**Proposition 16.1.4 — Mixed Product Rule.**

$$(A \otimes B)(C \otimes D) = AC \otimes BD \quad (16.9)$$

and,

$$(A \otimes B)^p = A^p \otimes B^p, \quad \forall p \in \mathbb{N} \quad (16.10)$$

**Proposition 16.1.5**

$$(A \otimes B)^T = A^T \otimes B^T \quad (16.11)$$

**Proposition 16.1.6**

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \quad (16.12)$$

**Proposition 16.1.7**

$$\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B) \quad (16.13)$$

**Proposition 16.1.8**

$$\text{tr}(A \otimes B) = \text{tr}(B \otimes A) = \text{tr}(A)\text{tr}(B) \quad (16.14)$$

**Proposition 16.1.9** Let  $A \in \mathcal{M}_{n \times n}$  and  $B \in \mathcal{M}_{m \times m}$ , we have,

$$\text{mspec}(A \otimes B) = \{\lambda \mu, \lambda \in \text{mspec}(A), \mu \in \text{mspec}(B)\} \quad (16.15)$$

**Proposition 16.1.10** Let  $A \in \mathcal{M}_{n \times n}$  and  $B \in \mathcal{M}_{m \times m}$ , we have,

$$\det(A \otimes B) = (\det A)^m (\det B)^n \quad (16.16)$$

We deduce from 16.15,

**Proposition 16.1.11** Let  $A \in \mathcal{M}_{n \times n}$ , we have,

$$\rho(A \otimes A) = \rho(A)^2 \quad (16.17)$$

**Proposition 16.1.12** Let  $f$  be an analytic function,  $A \in \mathcal{M}_{n \times n}$  such that  $f(A)$  exists, then we have,

$$f(I_m \otimes A) = I_m \otimes f(A) \quad (16.18)$$

$$f(A \otimes I_m) = f(A) \otimes I_m \quad (16.19)$$

**Proposition 16.1.13** Let  $X \in \mathbb{R}^n$  and  $Y \in \mathbb{R}^m$ , be two vectors. We have,

$$XY^T = X \otimes (Y^T) = (Y^T) \otimes X \quad (16.20)$$

moreover, we have,

$$\text{vec}(XY^T) = Y \otimes X \quad (16.21)$$

**Definition 16.1.3 — Kronecker permutation matrix.** The Kronecker permutation matrix  $P_{n,m} \in \mathcal{M}_{nm \times nm}$ , is defined by,

$$P_{n,m} = \sum_{i,j=1}^{n,m} E_{i,j,n \times m} \otimes E_{j,i,m \times n} \quad (16.22)$$

**Proposition 16.1.14** Let  $A \in \mathcal{M}_{m \times n}$ , we have,

$$\text{vec}(A^T) = P_{m,n} \text{vec}(A) \quad (16.23)$$

**Proposition 16.1.15** Let us consider the Kronecker permutation matrix  $P_{n,m} \in \mathcal{M}_{nm \times nm}$ . Then we have,

- $P_{n,m}^T = P_{n,m}^{-1} = P_{m,n}$
- $P_{n,m}$  is orthogonal,
- $P_{n,m} P_{m,n} = I_{nm}$
- $P_{n,n}$  is orthogonal, symmetric and involutory,
- $P_{n,n}$  is a reflector,
- $\text{tr} P_{n,n} = n$ ,
- $P_{1,m} = I_m$ , and  $P_{n,1} = I_n$
- if  $X \in \mathbb{R}^n$  and  $Y \in \mathbb{R}^m$ , then,

$$P_{n,m}(Y \otimes X) = X \otimes Y \quad (16.24)$$

- if  $A \in \mathcal{M}_{n \times m}$  and  $B \in \mathcal{M}_{r \times s}$ , then

$$P_{r,n}(A \otimes B)P_{m,s} = B \otimes A \quad (16.25)$$

- if  $A \in \mathcal{M}_{n \times n}$  and  $B \in \mathcal{M}_{m \times m}$ , then

$$P_{m,n}(A \otimes B)P_{n,m} = P_{m,n}(A \otimes B)P_{m,n}^{-1} = B \otimes A \quad (16.26)$$

Therefore,  $A \otimes B$  and  $B \otimes A$  are similar.



**Proposition 16.1.16** Let  $A \in \mathcal{M}_{n \times n}$  and  $B \in \mathcal{M}_{m \times m}$ , then we have the following properties,

- if  $A$  and  $B$  are diagonal, then  $A \otimes B$  is diagonal,
- if  $A$  and  $B$  are upper triangular, then  $A \otimes B$  is upper triangular,
- if  $A$  and  $B$  are lower triangular, then  $A \otimes B$  is lower triangular,

**Proposition 16.1.17** Let  $A, C \in \mathcal{M}_{n \times m}$  and  $B, D \in \mathcal{M}_{r \times s}$ . If  $A$  is (left equivalent, right equivalent, equivalent) to  $C$ , and assume that  $B$  is (left equivalent, right equivalent, equivalent) to  $D$ . Then,  $A \otimes B$  is (left equivalent, right equivalent, equivalent) to  $C \otimes D$ .

**R** The use of Kronecker product preconditioners is well known [Langville\_Stewart, Elisabeth\_Ullmann, GRIGORI:2008:INRIA-00268301:5, 14], it is based on results of the form,

$$\text{Minimizing, } \phi_A(B, C) = \|A - B \otimes C\|^2 \quad (16.27)$$

for a chosen norm.

### 16.1.1 Kronecker sum

**Definition 16.1.4 — Kronecker sum.** Let  $A = (a_{ij}) \in \mathcal{M}_{n \times n}$  and  $B = (b_{ij}) \in \mathcal{M}_{m \times m}$  be two matrices. The Kronecker sum of  $A$  and  $B$ , denoted by  $A \oplus B \in \mathcal{M}_{mn \times mn}$ , defines the following matrix:

$$A \oplus B = A \otimes I_m + I_n \otimes B \quad (16.28)$$

**Proposition 16.1.18** Let  $A \in \mathcal{M}_{n \times n}$  and  $B \in \mathcal{M}_{m \times m}$ , we have,

$$\text{mspec}(A \oplus B) = \{\lambda + \mu, \lambda \in \text{mspec}(A), \mu \in \text{mspec}(B)\} \quad (16.29)$$

### 16.1.2 Solving $AX + XB = C$

Let  $A \in \mathcal{M}_{n \times n}$ ,  $B \in \mathcal{M}_{m \times m}$  and  $C \in \mathcal{M}_{n \times m}$ . The aim of this section, is to solve the equation:

$$AX + XB = C \quad (16.30)$$

we can rewrite this equation in term of the Kronecker sum:

$$(B^T \oplus A) \text{vec}X = \text{vec}C \quad (16.31)$$

or equivalently,

$$Gx = c \quad (16.32)$$

where,  $G = (B^T \oplus A)$ ,  $x = \text{vec}X$ , and  $c = \text{vec}C$ .

Using the property 16.29, we can easily check that 16.30 has a unique solution if and only if  $G$  is nonsingular, i.e  $\lambda + \mu \neq 0$ ,  $\forall \lambda \in \text{mspec}(A)$ ,  $\forall \mu \in \text{mspec}(B)$ , which can be written in the form,

$$\text{mspec}(A) \cap \text{mspec}(-B) = \emptyset \quad (16.33)$$

**Proposition 16.1.19** If  $\text{mspec}(A) \cap \text{mspec}(-B) = \emptyset$ , then there exists a unique matrix  $X \in \mathcal{M}_{n \times m}$ , satisfying 16.30. Moreover, the matrices  $\begin{pmatrix} A & C \\ 0 & -B \end{pmatrix}$  and  $\begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix}$  are similar and verify,

$$\begin{pmatrix} A & C \\ 0 & -B \end{pmatrix} = \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix}. \quad (16.34)$$

**16.1.3 Solving  $AXB = C$** 

Let  $A, B, C$  and  $X \in \mathcal{M}_{n \times n}$ . As seen previously, using 16.13, the equation

$$AXB = C \tag{16.35}$$

can be written in the form,

$$Hx = c \tag{16.36}$$

where,  $H = (B^T \otimes A)$ ,  $x = \text{vec}X$ , and  $c = \text{vec}C$ .

Using the property 16.15, we can easily check that 16.35 has a unique solution if and only if  $H$  is nonsingular, *i.e.*  $\lambda\mu \neq 0$ ,  $\forall \lambda \in \text{mspec}(A)$ ,  $\forall \mu \in \text{mspec}(B)$ , which is equivalent to,  $A$  and  $B$  are both nonsingular.

**16.1.4 Solving  $\sum_{i=1}^r A_i X B_i = C$** 

Let  $A_i, B_i, C$ ,  $1 \leq i \leq r$  and  $X \in \mathcal{M}_{n \times n}$ . Using, the previous result, it is easy to show that the solution of:

$$\sum_{i=1}^r A_i X B_i = C \tag{16.37}$$

can be written in the form,

$$Hx = c \tag{16.38}$$

where,  $H = \sum_{i=1}^r (B_i^T \otimes A_i)$ ,  $x = \text{vec}X$ , and  $c = \text{vec}C$ .

**16.2 Problems**

TODO

## 17. Historical Notes

TODO

# VI

# Applications

<b>18</b>	<b>Navier Stokes</b> .....	<b>150</b>
18.1	Problems	
<b>19</b>	<b>Historical Notes</b> .....	<b>151</b>
	<b>Bibliography</b> .....	<b>152</b>
	Articles	
	Books	



# 18. Navier Stokes

## 18.1 Problems

TODO

## 19. Historical Notes

TODO

## Bibliography

### Articles

- [2] Sylvie Boldo and Marc Daumas. “A Simple Test Qualifying the Accuracy of Horner’S Rule for Polynomials”. In: *Numerical Algorithms* 37 (2004), pages 45–60 (cited on page 13).
- [7] R.T. Farouki and V.T. Rajan. “On the numerical condition of polynomials in Bernstein form”. In: *Computer Aided Geometric Design* 4.3 (1987), pages 191–216. ISSN: 0167-8396. DOI: [https://doi.org/10.1016/0167-8396\(87\)90012-4](https://doi.org/10.1016/0167-8396(87)90012-4). URL: <https://www.sciencedirect.com/science/article/pii/0167839687900124> (cited on page 13).
- [8] R.T. Farouki and V.T. Rajan. “Algorithms for polynomials in Bernstein form”. In: *Computer Aided Geometric Design* 5.1 (1988), pages 1–26. ISSN: 0167-8396. DOI: [https://doi.org/10.1016/0167-8396\(88\)90016-7](https://doi.org/10.1016/0167-8396(88)90016-7). URL: <https://www.sciencedirect.com/science/article/pii/0167839688900167> (cited on page 13).
- [13] Qixing Huang, Shimin Hu, and Ralph Robert Martin. “Fast degree elevation and knot insertion for B-spline curves”. In: *Comput. Aided Geom. Des.* 22 (2005), pages 183–197 (cited on pages 62, 63).
- [14] Charles F. van Loan. “The ubiquitous Kronecker product”. In: *J. Comput. Appl. Math.* 123 (1-2 Nov. 2000), pages 85–100. ISSN: 0377-0427. DOI: 10.1016/S0377-0427(00)00393-9. URL: <http://portal.acm.org/citation.cfm?id=363882.363888> (cited on page 145).
- [18] Hartmut Prautzsch and Bruce Piper. “A fast algorithm to raise the degree of spline curves”. In: *Comput. Aided Geom. Des.* 8 (4 Oct. 1991), pages 253–265. ISSN: 0167-8396. DOI: 10.1016/0167-8396(91)90015-4. URL: <http://portal.acm.org/citation.cfm?id=124930.124932> (cited on page 63).



## Books

- [1] D. Boffi, F. Brezzi, and M. Fortin. *Mixed Finite Element Methods and Applications*. Springer Series in Computational Mathematics. Springer Berlin Heidelberg, 2013. ISBN: 9783642365195. URL: <https://books.google.co.ma/books?id=mRhAAAAAQBAJ> (cited on page 96).
- [3] C. de Boor. *A Practical Guide to Splines*. Applied Mathematical Sciences. Springer New York, 2001. ISBN: 9780387953663. URL: [https://books.google.de/books?id=m0QDJvBI%5C\\_ecC](https://books.google.de/books?id=m0QDJvBI%5C_ecC) (cited on page 66).
- [4] Elaine Cohen, Richard F Riesenfeld, and Gershon Elber. *Geometric modeling with splines: an introduction*. CRC Press, 2001 (cited on page 66).
- [5] Gerald Farin. *NURBS for Curve & Surface Design: From Projective Geometry to Practical Use*. CRC Press, 1999 (cited on page 66).
- [6] Gerald E Farin and Gerald Farin. *Curves and surfaces for CAGD: a practical guide*. Morgan Kaufmann, 2002 (cited on page 66).
- [9] V. Girault and P.A. Raviart. *Finite element methods for Navier-Stokes equations: theory and algorithms*. Springer series in computational mathematics. Springer-Verlag, 1986. ISBN: 9783540157960. URL: <https://books.google.de/books?id=BVfvAAAAMAAJ> (cited on page 99).
- [10] Ronald N. Goldman and Tom Lyche. *Knot Insertion and Deletion Algorithms for B-Spline Curves and Surfaces*. Edited by Ronald N. Goldman and Tom Lyche. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1992. DOI: 10.1137/1.9781611971583. eprint: <https://epubs.siam.org/doi/pdf/10.1137/1.9781611971583>. URL: <https://epubs.siam.org/doi/abs/10.1137/1.9781611971583> (cited on page 61).
- [11] A. Graham. *Kronecker products and matrix calculus with applications*. 1981 (cited on page 142).
- [12] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. Second. Society for Industrial and Applied Mathematics, 2002. DOI: 10.1137/1.9780898718027. eprint: <https://epubs.siam.org/doi/pdf/10.1137/1.9780898718027>. URL: <https://epubs.siam.org/doi/abs/10.1137/1.9780898718027> (cited on page 13).
- [15] Monk P. *Finite Element Methods for Maxwell's Equations*. Calderon Press (Oxford), 2003. ISBN: 978-0-19-850888-5 (cited on page 99).
- [16] Les Piegl and Wayne Tiller. *The NURBS book*. Springer Science & Business Media, 1996 (cited on pages 63, 66).
- [17] Hartmut Prautzsch, Wolfgang Boehm, and Marco Paluszny. *Bézier and B-spline techniques*. Springer Science & Business Media, 2002 (cited on page 66).
- [19] David F Rogers. *An introduction to NURBS: with historical perspective*. Morgan Kaufmann, 2001 (cited on page 66).
- [20] Bernstein D. S. *Matrix Mathematics: Theory, Facts, and Formulas*. second. PRINCETON UNIVERSITY PRESS, 2009. ISBN: 978-0-691-13287-7 (cited on page 142).