

# **Les éléments finis : de la théorie à la pratique**

André Fortin  
Professeur titulaire  
Département de mathématiques et de statistique  
Université Laval

et

André Garon  
Professeur titulaire  
Département de génie mécanique  
École Polytechnique de Montréal

©1997-2011



# Avant-propos

La résolution des équations différentielles ou plus généralement des équations aux dérivées partielles occupe une place importante en ingénierie et en mathématiques appliquées. Chacune de ces disciplines apporte une contribution différente mais complémentaire à la compréhension et à la résolution de tels problèmes.

Il existe plusieurs techniques permettant de résoudre les équations aux dérivées partielles. On pense par exemple aux méthodes de différences finies, de volumes finis, aux méthodes spectrales, etc. On peut sans aucun doute affirmer que la plus largement répandue est la méthode des éléments finis. Cette popularité n'est pas sans fondement. La méthode des éléments finis est très générale et possède une base mathématique rigoureuse qui est fort utile, même sur le plan très pratique. En effet, cette base mathématique permet de prévoir jusqu'à un certain point la précision de notre approximation et même d'améliorer cette précision, via les méthodes adaptatives.

Ce texte est donc une introduction à la méthode des éléments finis. Nous poursuivrons ainsi deux objectifs. Bien sûr, nous souhaitons introduire la méthode des éléments finis et en donner une description relativement classique. Mais notre principal objectif est d'en dégager aussi les bases mathématiques plus fondamentales. On peut se demander s'il y a vraiment besoin de s'attarder autant sur les aspects plus mathématiques. La réponse nous est apparue de plus en plus évidente au fur et à mesure que se développaient les multiples applications de cette méthode. Les notions de convergence, de normes, d'espaces fonctionnels sont de plus en plus nécessaires pour aborder les problèmes modernes notamment en ce qui concerne les méthodes adaptatives, les méthodes de stabilisation et le développement de discrétisations compatibles dans le cas de problèmes à plusieurs variables comme les équations de Navier-Stokes ou les problèmes de coques. Pour travailler sérieusement sur ces problèmes, une connaissance superficielle de la méthode des éléments finis ne suffit plus et on doit aller plus en profondeur.

Il va de soi que poursuivre ces deux objectifs ne va pas sans difficultés. Au risque de déplaire à tous, nous visons un auditoire assez vaste allant du débutant au lecteur plus aguerri ayant déjà une connaissance de base en éléments finis.

Cet ouvrage s'adresse donc principalement aux étudiants en ingénierie, bien que les étudiants en mathématiques pourront y voir un complément pratique à leur formation plus théorique. Nous implorons la patience des étudiants ingénieurs. Les premiers chapitres vous paraîtront peut-être très théoriques mais soyez assurés que nous avons réduit au minimum les considérations théoriques et que nous nous limitons à l'essentiel. Nous implorons aussi l'indulgence des lecteurs ayant une formation

mathématique plus avancée car, comme nous l'avons déjà mentionné, la rigueur mathématique n'est pas notre obsession, bien que nous ayons fait notre possible pour rester rigoureux.

Dans la mesure du possible, cet ouvrage est auto-suffisant. On retrouve en annexe quelques rappels de notions mathématiques élémentaires portant sur les tenseurs et les changements de systèmes de coordonnées qui sont d'une grande utilité dans l'étude des équations aux dérivées partielles. Une connaissance des méthodes d'analyse numérique élémentaire est requise et en particulier des notions d'interpolation de Lagrange et d'intégration numérique de Gauss qui sont également rappelées en annexe.

Enfin, nous souhaitons remercier tous ceux qui ont contribué, de près ou de loin, à la réalisation de cet ouvrage. De nombreux étudiants ont émis des commentaires constructifs qui nous ont incités à améliorer certains passages plus difficiles. Soyez tous assurés de notre reconnaissance.

# Table des matières

<b>1</b>	<b>Introduction et exemples</b>	<b>1</b>
1.1	Bref historique . . . . .	1
1.2	Applications . . . . .	1
1.3	Brève introduction à la problématique . . . . .	2
<b>2</b>	<b>Espaces fonctionnels</b>	<b>7</b>
2.1	Les distributions . . . . .	7
2.1.1	Définitions et propriétés générales . . . . .	7
2.1.2	Distributions de plusieurs variables . . . . .	20
2.2	Espaces fonctionnels linéaires . . . . .	24
2.3	Quelques espaces de Sobolev . . . . .	30
2.3.1	L'espace $H^1(\Omega)$ . . . . .	30
2.3.2	L'espace $H^2(\Omega)$ . . . . .	36
2.4	Un résultat d'approximation . . . . .	39
2.5	Exercices . . . . .	42
<b>3</b>	<b>Théorème de Lax-Milgram</b>	<b>43</b>
3.1	Formes linéaires et bilinéaires . . . . .	43
3.2	Théorème de Lax-Milgram . . . . .	46
3.3	Applications . . . . .	48
3.3.1	Problèmes d'ordre 2 . . . . .	48
3.3.2	Problèmes d'ordre 4 . . . . .	54
3.3.3	Résumé . . . . .	57
3.4	Exercices . . . . .	59
<b>4</b>	<b>Méthode de Ritz</b>	<b>63</b>
4.1	Principes généraux . . . . .	63
4.2	Exemples . . . . .	65
4.3	Exercices . . . . .	73

<b>5</b>	<b>Éléments finis unidimensionnels</b>	<b>77</b>
5.1	Équations différentielles d'ordre 2	77
5.1.1	Problème type	77
5.1.2	Le maillage	79
5.1.3	Formulation variationnelle élémentaire	85
5.1.4	Passage à l'élément de référence	87
5.1.5	Construction des fonctions d'interpolation $\hat{\psi}_i(\xi)$	89
5.1.6	Évaluation du système élémentaire	94
5.1.7	Assemblage	97
5.1.8	Imposition des conditions aux limites	107
5.1.9	Solution du système global	108
5.1.10	Présentation des résultats	111
5.1.11	Exemples et applications	111
5.2	Équations différentielles d'ordre 4	118
5.2.1	Problème type	118
5.2.2	Le maillage	120
5.2.3	Formulation variationnelle élémentaire	122
5.2.4	Passage à l'élément de référence	123
5.2.5	Construction des fonctions d'interpolation	124
5.2.6	Évaluation du système élémentaire	127
5.2.7	Assemblage	127
5.2.8	Imposition des conditions aux limites	128
5.2.9	Solution du système global	128
5.2.10	Exemples et applications	128
5.3	Exercices	135
<b>6</b>	<b>Éléments finis multidimensionnels</b>	<b>137</b>
6.1	Problème type	137
6.2	Le maillage	138
6.2.1	Les noeuds	138
6.2.2	Les degrés de liberté	139
6.2.3	Numérotation des degrés de liberté	139
6.3	Formulation variationnelle élémentaire	140
6.4	Passage à l'élément de référence	142
6.5	Construction des fonctions d'interpolation	151
6.6	Évaluation du système élémentaire	151
6.7	Assemblage	152
6.8	Imposition des conditions aux limites	152
6.9	Résolution du système linéaire global	153
6.10	Présentation des résultats	153
6.11	Exemples et applications	153

6.12 Exercices . . . . .	164
<b>7 Analyse de convergence</b>	<b>167</b>
7.1 Bases théoriques . . . . .	167
7.2 Quelques exemples . . . . .	170
<b>8 Problèmes non linéaires</b>	<b>181</b>
8.1 Rappel sur les systèmes d'équations non linéaires . . . . .	182
8.2 Dérivée d'une fonctionnelle . . . . .	186
8.3 Application aux formulations variationnelles . . . . .	187
8.4 Exercices . . . . .	193
<b>9 Problèmes instationnaires</b>	<b>195</b>
9.1 Rappel sur les équations différentielles ordinaires . . . . .	195
9.2 Formulation quasi-variationnelle . . . . .	197
9.3 Le theta-schéma . . . . .	199
9.3.1 Cas linéaire . . . . .	200
9.3.2 Cas linéaire où la matrice masse est constante . . . . .	200
9.3.3 Cas général non linéaire . . . . .	201
9.4 Un schéma implicite d'ordre 2 . . . . .	204
9.4.1 Variante de type points fixes . . . . .	205
9.5 Résultats numériques . . . . .	205
9.5.1 Problème thermique . . . . .	205
9.5.2 Croissance de populations . . . . .	208
<b>10 Problème de convection-diffusion et stabilisation SUPG</b>	<b>209</b>
10.1 Introduction . . . . .	209
10.1.1 Une première approche par éléments finis . . . . .	212
10.1.2 Une deuxième approche par éléments finis . . . . .	215
10.2 Exercices . . . . .	218
<b>11 Application aux problèmes d'élasticité</b>	<b>219</b>
11.1 Problèmes d'élasticité linéaire . . . . .	219
11.2 Matériau linéaire élastique isotrope . . . . .	223
11.3 Matériau linéaire élastique orthotrope . . . . .	225
11.4 État plan de contraintes et état plan de déformation . . . . .	226
11.4.1 État plan de déformation . . . . .	227
11.4.2 État plan de contraintes . . . . .	227
11.5 Le maillage . . . . .	228
11.6 Formulation variationnelle élémentaire . . . . .	231
11.7 Construction des fonctions d'interpolation . . . . .	232
11.8 Passage à l'élément de référence . . . . .	234

11.9	Évaluation du système élémentaire . . . . .	235
11.10	Assemblage et imposition des conditions aux limites . . . . .	236
11.11	Résolution du système global . . . . .	236
11.12	Visualisation des résultats . . . . .	236
11.13	Applications . . . . .	236
11.13.1	Essai en cisaillement . . . . .	236
11.13.2	Plaque trouée en élongation . . . . .	236
11.14	Exercices . . . . .	242
<b>12</b>	<b>Problème de Stokes</b> . . . . .	<b>243</b>
12.1	Introduction . . . . .	243
12.2	Le problème continu . . . . .	244
12.2.1	Cadre fonctionnel . . . . .	246
12.2.2	Existence et unicité . . . . .	247
12.3	Le problème discret . . . . .	248
12.3.1	Quelques résultats théoriques importants . . . . .	249
12.4	Formulation point-selle (facultative) . . . . .	250
12.5	Formulation variationnelle élémentaire . . . . .	251
12.6	Choix des éléments . . . . .	253
12.6.1	Pression continue . . . . .	253
12.6.2	Pression discontinue . . . . .	256
12.7	Résultats numériques : cas newtonien . . . . .	258
12.7.1	L'écoulement de Poiseuille . . . . .	258
12.8	Problème de Stokes non linéaire . . . . .	259
12.8.1	Résultats numériques : cas viscoplastique . . . . .	263
12.9	Les équations de Navier-Stokes . . . . .	266
12.9.1	Résultats numériques : cas Navier-Stokes . . . . .	269
12.10	Exercices . . . . .	274
<b>13</b>	<b>Formulations mixtes en élasticité linéaire</b> . . . . .	<b>275</b>
13.1	Cas isotrope . . . . .	275
13.2	Cas anisotrope . . . . .	279
13.3	Résultats numériques . . . . .	281
13.3.1	Plaque trouée en élongation . . . . .	281
13.4	Exercices . . . . .	287
<b>14</b>	<b>Matériaux en grandes déformations</b> . . . . .	<b>289</b>
14.1	Introduction . . . . .	289
14.2	Déformations et tenseurs associés . . . . .	289
14.3	Tenseurs de Green-Lagrange et Piola-Kirchoff . . . . .	295
14.4	Matériaux hyperélastiques . . . . .	299

14.4.1	Limite incompressible . . . . .	302
14.5	Formulations variationnelles . . . . .	305
14.5.1	Intégrales volumiques . . . . .	306
14.5.2	Intégrales surfaciques . . . . .	307
14.5.3	Pression suiveuse . . . . .	307
14.5.4	Formulation variationnelle sur la configuration initiale . . . . .	308
14.6	Linéarisation et méthode de Newton . . . . .	309
14.6.1	Formulation en déplacement seulement . . . . .	310
14.6.2	Formulation mixte en déplacements-pression . . . . .	312
14.6.3	Formulation pénalisée . . . . .	315
14.7	Liens avec la géométrie différentielle intrinsèque . . . . .	316
14.7.1	Variation du jacobien . . . . .	316
14.7.2	Variation du jacobien surfacique . . . . .	316
14.7.3	Variation de la normale . . . . .	318
14.7.4	Pression suiveuse . . . . .	319
14.8	Formulation Lagrangienne actualisée . . . . .	320
14.8.1	Formulation en déplacement seulement . . . . .	321
14.8.2	Formulation Lagrangienne actualisée mixte . . . . .	325
14.8.3	Équivalence des formulations . . . . .	325
14.8.4	Algorithme complet . . . . .	325
14.9	Exercices . . . . .	327
<b>15</b>	<b>Problèmes de contact unilatéral et frottant</b>	<b>329</b>
15.1	Introduction . . . . .	329
15.1.1	Distance signée . . . . .	329
15.2	Contact unilatéral en petites déformations . . . . .	332
15.3	Contact unilatéral en grandes déformations . . . . .	334
15.3.1	Méthode de pénalisation du contact unilatéral . . . . .	335
15.3.2	Contact frottant . . . . .	336
<b>A</b>	<b>Rappels sur le théorème de la divergence</b>	<b>341</b>
A.1	Gradient, divergence et laplacien . . . . .	341
A.2	Intégrales curvilignes et surfaciques . . . . .	342
A.2.1	Rappel sur les intégrales curvilignes . . . . .	342
A.2.2	Rappel sur les intégrales surfaciques . . . . .	342
A.3	Théorème de la divergence . . . . .	343
A.4	Transformation de Piola . . . . .	347
A.5	Exercices . . . . .	349

<b>B</b>	<b>Rappels sur les tenseurs</b>	<b>351</b>
B.1	Notions de base . . . . .	351
B.1.1	Les vecteurs . . . . .	351
B.1.2	Les tenseurs d'ordre 2 . . . . .	353
B.2	Calcul variationnel avec les tenseurs . . . . .	356
B.2.1	Résultats généraux . . . . .	356
B.2.2	Dérivées des invariants . . . . .	359
B.2.3	Application aux grandes déformations . . . . .	361
B.3	Exercices . . . . .	364
<b>C</b>	<b>Interpolation de Lagrange</b>	<b>365</b>
C.1	Interpolation en dimension 1 . . . . .	366
C.2	Interpolation en dimension 2 . . . . .	368
C.2.1	Interpolation sur les triangles . . . . .	369
C.2.2	Sur les quadrilatères . . . . .	372
C.3	Interpolation en dimension 3 . . . . .	374
C.3.1	Sur les tétraèdres . . . . .	374
<b>D</b>	<b>Intégration numérique</b>	<b>377</b>
D.1	En dimension 1 . . . . .	377
D.2	En dimension 2 ou 3 . . . . .	380
D.2.1	Sur les quadrilatères . . . . .	380
D.2.2	Sur les triangles . . . . .	381
D.2.3	Sur les tétraèdres . . . . .	381
	Réponses aux exercices du chapitre 2 . . . . .	385
	Réponses aux exercices du chapitre 3 . . . . .	386
	Réponses aux exercices du chapitre 4 . . . . .	388
	Réponses aux exercices du chapitre 5 . . . . .	390
	Réponses aux exercices du chapitre 6 . . . . .	391
	Réponses aux exercices du chapitre 8 . . . . .	392
	Réponses aux exercices du chapitre 14 . . . . .	394

# Table des figures

1.1	Solution par différences finies . . . . .	3
2.1	Fonction $\phi(x) = e^{\frac{R^2}{(x-p)^2 - R^2}}$ ( $R = 2$ , $p = 0$ ) . . . . .	10
2.2	Fonction $\phi(x) = e^{\frac{R^2}{x_1^2 + x_2^2 - R^2}}$ ( $R = 2$ et $\mathbf{p} = 0$ ) . . . . .	10
2.3	Espaces de fonctions $L^1_{loc}(\Omega)$ et de distributions $\mathcal{L}^1_{loc}(\Omega)$ . . . . .	15
2.4	Fonction de Heaviside . . . . .	17
2.5	Fonction $f(x) =  x $ . . . . .	18
2.6	Domaines $\Omega_1$ et $\Omega_2$ séparés par une courbe $C$ . . . . .	22
2.7	Espace $L^2(\Omega)$ . . . . .	27
4.1	Solutions $u(x)$ et $u'(x)$ pour $N = 1$ , $N = 3$ et $N = 5$ . . . . .	67
4.2	Méthode de Ritz ( $N = 2$ ) . . . . .	69
4.3	Géométrie et conditions aux limites . . . . .	69
4.4	Méthode de Ritz en dimension 2 : fonctions $\delta_u(x)$ et $u(x)$ . . . . .	71
5.1	Maillage en dimension 1 . . . . .	79
5.2	Noeuds géométriques et de calcul de l'élément $K$ . . . . .	80
5.3	Maillage : $n_{el} = 3, n_g^K = 2, n_c^K = 3$ . . . . .	82
5.4	Fonctions d'interpolation linéaires sur $\hat{K}$ . . . . .	91
5.5	Fonctions d'interpolation linéaires sur $K$ . . . . .	92
5.6	Fonctions d'interpolation quadratiques sur $\hat{K}$ . . . . .	93
5.7	Fonctions d'interpolation quadratiques sur $K$ . . . . .	94
5.8	Fonctions d'interpolation hiérarchiques sur $K$ . . . . .	95
5.9	Fonctions de Ritz linéaires par élément . . . . .	99
5.10	Fonctions de Ritz quadratiques par élément . . . . .	101
5.11	Déflexion verticale d'un câble . . . . .	112
5.12	Maillage pour la déflexion verticale d'un câble . . . . .	113
5.13	Solution numérique : $u(x)$ et $T(x)u'(x)$ . . . . .	119
5.14	Fonctions d'interpolation d'Hermite ( $h^K = 1$ ) . . . . .	126
5.15	Problème de la poutre encastrée . . . . .	129

5.16	Déformation de la poutre encastrée . . . . .	134
6.1	Structure de la matrice avant et après renumérotation . . . . .	141
6.2	Élément de référence triangulaire à 3 noeuds ( $n_g^K = 3$ ) . . . . .	142
6.3	Élément de référence triangulaire à 6 noeuds ( $n_g^K = 6$ ) . . . . .	143
6.4	Élément de référence quadrangulaire à 4 noeuds ( $n_g^K = 4$ ) . . . . .	143
6.5	Élément de référence quadrangulaire à 9 noeuds ( $n_g^K = 9$ ) . . . . .	143
6.6	Élément de référence tétraédrique à 4 noeuds ( $n_g^K = 4$ ) . . . . .	144
6.7	Transformation linéaire sur un triangle ( $n_g^K = 3$ ) . . . . .	146
6.8	Transformation quadratique sur un triangle ( $n_g^K = 6$ ) . . . . .	149
6.9	Transformation linéaire sur un tétraèdre ( $n_g^K = 4$ ) . . . . .	150
6.10	Numérotation des éléments et des noeuds . . . . .	154
6.11	Isovaleurs de la fonction $u(x)$ et vue tridimensionnelle . . . . .	161
6.12	Géométrie et conditions aux limites . . . . .	161
6.13	Transfert thermique . . . . .	162
6.14	Élément de référence triangulaire de degré 3 ( $n_g^K = 10$ ) . . . . .	164
6.15	Maillage et numéros de degrés de liberté . . . . .	165
6.16	Élément $K$ . . . . .	166
7.1	Maillage de $32 \times 32$ éléments . . . . .	173
7.2	Erreurs commises pour les approximations linéaire et quadratique . . . . .	175
7.3	Transfert thermique : maillages utilisés . . . . .	177
7.4	Transfert thermique : isothermes . . . . .	178
7.5	Coupe de la température sur l'axe $y = 1$ . . . . .	179
9.1	Transfert thermique instationnaire . . . . .	206
9.2	Transfert thermique instationnaire (suite) . . . . .	207
10.1	Maillage de deux éléments . . . . .	210
10.2	Équation de convection-diffusion : différences centrées et éléments finis linéaires . . . . .	211
10.3	Équation de convection-diffusion : différences arrières . . . . .	213
10.4	Fonction de Ritz linéaire de type Galerkin . . . . .	214
10.5	Fonction de Ritz modifiée de type SUPG . . . . .	215
10.6	Schéma de type SUPG avec $\alpha$ défini par 10.6 . . . . .	216
10.7	Fonction de Ritz modifiée de type SUPG . . . . .	217
11.1	Maillage, numéros d'éléments et numéros noeuds de calcul . . . . .	229
11.2	Maillage utilisé . . . . .	237
11.3	Déformation du maillage . . . . .	238
11.4	Déplacement vertical . . . . .	238
11.5	Plaque trouée : géométrie et conditions aux limites . . . . .	239

11.6 Déplacements : $\nu = 0,3$ , $\nu = 0,4$ et $\nu = 0,4999$	240
11.7 Contrainte $\sigma_{11}$ : $\nu = 0,3$ , $\nu = 0,4$ et $\nu = 0,4999$	241
12.1 Élément Mini ( $O(h)$ )	254
12.2 Élément Mini en dimension 3 ( $O(h)$ )	254
12.3 Élément de Taylor-Hood $P_2 - P_1$ ( $O(h^2)$ )	255
12.4 Élément $P_2 - P_1$ en dimension 3 ( $O(h^2)$ )	255
12.5 Élément $Q_2 - Q_1$ ( $O(h^2)$ )	255
12.6 Élément $Q_2 - Q_1$ en dimension 3 ( $O(h^2)$ )	256
12.7 Élément incompatible à pression discontinue $Q_1 - Q_0$	256
12.8 Élément incompatible $Q_1 - Q_0$ en dimension 3	257
12.9 Élément à pression discontinue $P_2 - P_0$ ( $O(h)$ )	257
12.10 Élément à pression discontinue de Crouzeix-Raviart $P_2^+ - P_1$ ( $O(h^2)$ )	257
12.11 Élément à pression discontinue $Q_2 - P_1$ ( $O(h^2)$ )	257
12.12 Écoulement de Poiseuille (géométrie et conditions aux limites)	258
12.13 Écoulement de Poiseuille : maillages rectangulaire et triangulaire	260
12.14 Écoulement de Poiseuille pour l'élément $Q_1 - Q_0$ : vitesse et pression	260
12.15 Écoulement de Poiseuille pour l'élément Mini : vitesse et pression	261
12.16 Écoulement de Poiseuille pour l'élément $P_2 - P_1$ : vitesse et pression	261
12.17 Écoulement de Poiseuille : loi puissance avec $m = 1$ , $m = 0,5$ et $m = 0,3$	264
12.18 Profil de vitesse en sortie du canal	265
12.19 Perte de charge dans le canal	265
12.20 Maillage, perte de charge et module de vitesse en sortie de filière	267
12.21 Écoulement autour d'un cylindre : géométrie et conditions aux limites	270
12.22 Maillage et portrait global des lignes de courant à $Re = 1$	270
12.23 Lignes de courant derrière le cylindre à $Re = 1, 50$ et $100$	271
12.24 Champ de vitesse derrière le cylindre à $Re = 100$ sur une période	273
13.1 Déplacements : $\nu = 0,3$ , $\nu = 0,4$ et $\nu = 0,4999$	283
13.2 Pression : $\nu = 0,3$ , $\nu = 0,4$ et $\nu = 0,4999$	284
13.3 Contrainte $\sigma_{11}$ : $\nu = 0,3$ , $\nu = 0,4$ et $\nu = 0,4999$	285
13.4 Contrainte $\sigma_{11}$ et $p$ : $\nu = 0,4999999$	286
14.1 Description cinématique	290
14.2 Comparaison d'une condition de Neumann et d'une pression suiveuse	313
15.1 Fonction distance signée (par projection)	330
15.2 Lois de contact unilatéral originale (à gauche) et pénalisée (à droite)	336
15.3 Lois de frottement de Coulomb originale et régularisée	338
A.1 Ouvert $\Omega$ de frontière $\Gamma$	343
A.2 Transformation	346

A.3 Transformation d'un élément de longueur . . . . .	346
---	-----

# Chapitre 1

## Introduction et exemples

### 1.1 Bref historique

L'idée fondamentale derrière la méthode des éléments finis remonte loin en arrière. Les grecs par exemple avaient reconnu que l'on peut approcher la solution d'un problème complexe en le divisant en problèmes plus simples. On peut par exemple approcher le périmètre d'un cercle en calculant le périmètre d'un polygone à  $n$  côtés, pourvu que  $n$  soit suffisamment grand. Il suffit alors de connaître la longueur d'un segment de droite, problème beaucoup plus simple que celui de la longueur d'un arc de cercle.

L'application à la solution des équations aux dérivées partielles est évidemment plus récente et est intimement liée au développement de l'informatique. Courant [10] a introduit le concept de formulation variationnelle, qui est à la base de toute méthode d'éléments finis.

Pour la méthode de Ritz [32], on part d'un problème posé dans un espace de dimension infinie. On approche ensuite la solution du problème initial en cherchant une solution dans une suite croissante de sous-espaces de dimension finie. Ces problèmes approchés sont en général beaucoup plus facile à résoudre. On peut de plus espérer que la solution du problème en dimension infinie peut être obtenue par un passage à la limite. Le choix des fonctions de base constituant ces espaces de dimension finie est délicat et initialement on les construisait globalement sur le domaine. C'est Courant qui eut l'idée d'introduire des fonctions à support local qui simplifient grandement leur construction.

La théorie derrière la méthode des éléments finis a pris une forme plus rigoureuse avec les travaux de Strang et Fix [35]. à compléter ...

### 1.2 Applications

On retrouve les premières applications véritables de la méthode des éléments finis en 1956 en mécanique des structures. Un groupe de chercheurs (Turner, Clough, Martin et Topp [37]) de Boeing utilisent cette méthode pour calculer la voilure d'un avion. à compléter ...

La méthode des éléments finis est maintenant reconnue comme l'une des principales méthodes de résolution des équations aux dérivées partielles (EDP) dans des géométries quelconques, que ce soit en dimension un, deux ou trois. On trouve même des méthodes d'éléments finis en dimension 4, soit en espace-temps...

Les applications sont tout aussi nombreuses et variées. Les ingénieurs de diverses disciplines utilisent les éléments finis, que ce soit en mécanique des fluides ou des solides, mais aussi pour les problèmes thermiques, électro-magnétiques, chimiques, etc. On retrouve aussi des applications en physique, et notamment en astrophysique. à compléter ...

### 1.3 Brève introduction à la problématique

Pour illustrer le fonctionnement de la méthode des éléments finis et pour justifier l'introduction d'un certain nombre d'outils mathématiques, nous allons considérer un exemple très simple et effectuer une comparaison avec la méthode des différences finies.

Soit donc l'équation différentielle :

$$\begin{cases} -u''(x) = f(x) \\ u(0) = u(1) = 0 \end{cases} \quad (1.1)$$

où  $f(x)$  est une fonction connue. On cherche donc à obtenir une approximation de la solution  $u(x)$  dans l'intervalle  $[0, 1]$ . Pour ce faire, subdivisons cet intervalle en  $N$  sous-intervalles de longueur  $h = 1/N$  (les sous-intervalles peuvent éventuellement être de longueurs différentes). On obtient ainsi  $N + 1$  points  $x_i$  vérifiant  $x_0 = 0$ ,  $x_N = 1$  et pour les points intermédiaires :

$$x_i = \frac{i}{N}, \quad x_{i+1} - x_i = h$$

On note  $u_i$ , l'approximation de  $u(x_i)$  au point  $x_i$ . Les conditions aux limites imposent que  $u_0 = u_N = 0$ . La méthode des différences finies consiste à *discrétiser* directement l'équation différentielle en remplaçant les dérivées de  $u(x)$  par des *différences finies* et ce, en chaque point  $x_i$ . On peut par exemple utiliser une différence centrée d'ordre 2 (voir Fortin, réf. [17]) :

$$u''(x_i) = \frac{u(x_{i-1}) - 2u(x_i) + u(x_{i+1}))}{h^2} + O(h^2) \simeq \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2}$$

L'équation différentielle s'écrit en chaque point  $x_i$  :

$$-u''(x_i) = f(x_i)$$

de sorte qu'en remplaçant par la différence centrée, on obtient :

$$-\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} = f(x_i)$$

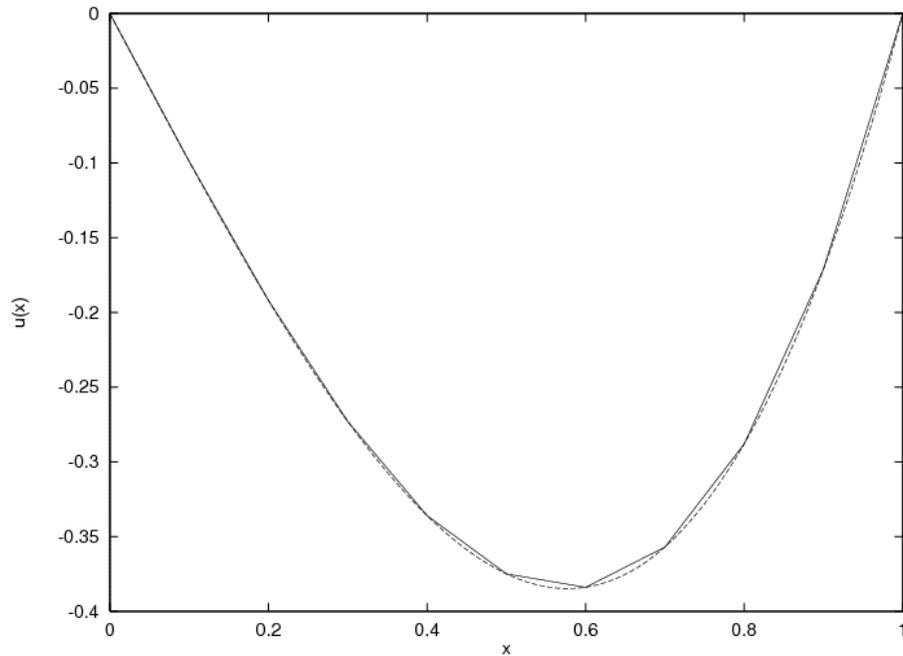


Figure 1.1 – Solution par différences finies

et ce pour  $i$  allant de 1 jusqu'à  $N - 1$ . Dans l'équation précédente, on a bien sûr négligé le terme d'erreur  $O(h^2)$  et il en résulte une approximation d'ordre 2. On obtient ainsi un système linéaire de  $(N - 1)$  équations en  $(N - 1)$  inconnues de la forme :

$$\begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 2 & -1 \\ 0 & 0 & \cdots & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{N-2} \\ u_{N-1} \end{bmatrix} = h^2 \begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \\ f(x_{N-2}) \\ f(x_{N-1}) \end{bmatrix}$$

La résolution de ce système linéaire tridiagonal est simple et fournit une solution approximative de l'équation différentielle de départ aux points  $x_i$ . Ainsi, à la figure 1.1, on peut voir la solution numérique obtenue avec 10 intervalles ( $h = 0,1$ ) et la fonction  $f(x) = -6x$ . Dans ce cas, on vérifie facilement que la solution analytique est  $u(x) = x^3 - x$ . On peut dès lors constater que la solution numérique est une bonne approximation de la solution analytique.

La méthode des éléments finis procède tout autrement. Nous allons donc reprendre l'exemple précédent par cette méthode, mais sans justifier toutes les étapes puisque nous ne sommes pas

encore en mesure de le faire. Multiplions l'équation différentielle 1.1 par une fonction dite *test*  $w(x)$  que nous supposons pour l'instant quelconque et intégrons sur l'intervalle  $[0, 1]$ . On obtient :

$$\int_0^1 -u''(x)w(x)dx = \int_0^1 f(x)w(x)dx$$

Intégrons maintenant par parties (en supposant que  $w(x)$  soit suffisamment régulière) pour obtenir :

$$\int_0^1 u'(x)w'(x)dx - (u'(1)w(1) - u'(0)w(0)) = \int_0^1 f(x)w(x)dx$$

Si on suppose maintenant que  $w(0) = w(1) = 0$ , on obtient une *formulation variationnelle* qui consiste à déterminer une fonction  $u(x)$ , vérifiant  $u(0) = u(1) = 0$  et telle que :

$$\int_0^1 u'(x)w'(x)dx = \int_0^1 f(x)w(x)dx \quad \forall w(x) \mid w(0) = w(1) = 0 \quad (1.2)$$

#### ◆ Remarque 1.1

La formulation variationnelle 1.2 ne fait apparaître que la dérivée d'ordre 1 de la fonction  $u(x)$  (et de la fonction test  $w(x)$ ) et demande donc moins de régularité que l'équation différentielle elle-même qui, dans ce cas, est d'ordre 2. La formulation variationnelle 1.2 est donc aussi appelée *formulation faible* (ou *forme faible*) de l'équation différentielle 1.1 par opposition à l'équation différentielle elle-même dite *formulation forte*. ◆

Il est évident que toute solution de 1.1 vérifie la formulation variationnelle 1.2. Par contre, l'inverse n'est pas aussi évident. Démontrons-le, toujours formellement. Soit donc  $u(x)$  une solution de la formulation variationnelle 1.2. Essayons de refaire le chemin inverse, jusqu'à l'équation différentielle 1.1. En intégrant par parties, on trouve :

$$\int_0^1 -u''(x)w(x)dx + u'(1)w(1) - u'(0)w(0) = \int_0^1 f(x)w(x)dx$$

c.-à-d.

$$\int_0^1 (-u''(x) - f(x)) w(x)dx = 0$$

qui est valide pour toute fonction  $w(x)$  qui s'annule en 0 et 1. Il nous faut déduire de cette équation que :

$$-u''(x) - f(x) = 0$$

Pour y arriver, il faut effectuer un choix judicieux de fonction  $w(x)$ . Considérons d'abord la fonction  $\phi(x) = x(1-x)$  qui est positive dans l'intervalle  $]0, 1[$  et qui s'annule en  $x = 0$  et  $x = 1$ . Choisissons ensuite :

$$w(x) = \phi(x)(-u''(x) - f(x))$$

Cette fonction test  $w(x)$  s'annule bien en  $x = 0$  et en  $x = 1$  et est donc admissible. On obtient ainsi :

$$\int_0^1 (-u''(x) - f(x))^2 \phi(x) dx = 0$$

Puisque l'intégrand est toujours positif et que son intégrale est nulle, on en conclut que :

$$(-u''(x) - f(x))^2 \phi(x) = 0$$

Puisque  $\phi(x)$  ne s'annule pas à l'intérieur de l'intervalle, on peut affirmer que :

$$-u''(x) = f(x) \quad \text{dans } ]0, 1[$$

et donc que  $u(x)$  vérifie l'équation différentielle. Il y a donc équivalence entre l'équation différentielle 1.1 et la formulation variationnelle 1.2.

Plusieurs questions importantes demeurent en ce qui concerne le développement précédent dont plusieurs étapes restent largement à justifier.

1. Comment déterminer précisément les fonctions tests  $w(x)$  admissibles ?
2. Sous quelles conditions la formulation variationnelle (1.2) possède-t-elle une solution unique ?
3. Sous quelles conditions la formulation variationnelle (*formulation faible*) est-elle équivalente à l'équation différentielle (*formulation forte*) ?
4. Comment peut-on discrétiser cette formulation variationnelle pour en tirer une solution numérique comme nous l'avons fait par la méthode des différences finies ?
5. Cette solution numérique converge-t-elle vers la solution analytique ? Et à quelle vitesse ?

C'est à toutes ces questions et à d'autres que nous tâcherons de répondre dans ce texte. La méthode des éléments finis est de fait une technique d'approximation des formulations variationnelles des équations aux dérivées partielles. En essayant de répondre aux questions précédentes, nous verrons la nécessité d'introduire un certain nombre d'outils mathématiques. C'est l'objet des deux chapitres suivants.



## Chapitre 2

# Espaces fonctionnels

L'analyse de la méthode des éléments finis requiert une bonne dose d'analyse fonctionnelle, outil fondamental pour une véritable compréhension de cette méthode. C'est l'objet de ce chapitre. Précisons dès le départ que notre objectif n'est pas de donner un cours d'analyse fonctionnelle complet mais bien de donner les outils de base nécessaires à l'utilisation efficace de la méthode des éléments finis.

Parmi les outils de base, on retrouve les notions de distributions, d'espaces de Hilbert, de Sobolev, etc. L'étude de ces notions pourrait faire l'objet d'un livre (et même de plusieurs) et il va de soi que nous nous contenterons d'un survol assez rapide mais relativement complet. Nous omettrons toutefois beaucoup de détails et de subtilités qui ont bien sûr leur importance mais qui ne sont pas essentielles à une bonne compréhension de la méthode des éléments finis.

### 2.1 Les distributions

Les distributions sont aux fonctions ce que les nombres irrationnels sont aux nombres rationnels. Les distributions sont en fait une généralisation de la notion de fonction. Nous en ferons une présentation sommaire en nous limitant aux notions essentielles comme la dérivation d'une distribution. Nous référons le lecteur à Gasquet-Witomski, réf. [21] pour un traitement simple et moderne de la théorie des distributions. Le lecteur plus averti peut consulter le texte de L. Schwartz, réf. [34] pour un traitement plus complet et plus classique.

#### 2.1.1 Définitions et propriétés générales

Dans ce qui suit,  $\Omega$  désignera un ensemble ouvert de  $R^n$  dont la frontière  $\Gamma$  est régulière. Rappelons maintenant deux notions importantes pour la suite.

**Définition 2.1**

Le *support* d'une fonction  $f(\mathbf{x})$  est le plus petit ensemble fermé de valeurs de  $\mathbf{x}$  en dehors duquel la fonction  $f(\mathbf{x})$  est identiquement nulle. C'est donc la fermeture de l'ensemble des points  $\mathbf{x}$  tels que  $f(\mathbf{x}) \neq 0$ .

**Définition 2.2**

Un sous-ensemble de  $\Omega \subset \mathbb{R}^n$  est dit *compact* s'il est fermé et borné.

**► Exemple 2.1**

La fonction :

$$f(x) = \begin{cases} |x| & \text{si } x \in ]-1, 1[ \\ 0 & \text{ailleurs} \end{cases}$$

a comme support l'intervalle fermé  $[-1, 1]$  qui est un ensemble compact. ◀

Un ensemble compact  $A$  contient donc tous ses points frontières et de plus, puisqu'il est borné, il existe une constante  $M$  telle que :

$$A \subset \{\mathbf{x} \mid \|\mathbf{x}\|_2 \leq M\}$$

où  $\|\mathbf{x}\|_2$  désigne la norme euclidienne du vecteur  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  :

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2} \quad (2.1)$$

Le compact  $A$  est ainsi inclus dans un disque de rayon  $M$ , pourvu que  $M$  soit suffisamment grand.

**Définition 2.3**

On appelle  $\mathcal{D}(\Omega)$  l'espace des fonctions infiniment différentiables sur  $\Omega$  et dont le support est compact et inclus dans  $\Omega$ .

Les fonctions de  $\mathcal{D}(\Omega)$  possèdent donc la propriété de s'annuler identiquement au voisinage du bord de  $\Omega$  ou lorsque la norme de  $\mathbf{x}$  est suffisamment grande, ce qui nous sera très utile par la suite. Elles sont de plus extrêmement régulières puisqu'infiniment différentiables et leurs dérivées s'annulent également au voisinage de la frontière du domaine  $\Omega$ .

► **Exemple 2.2**

Les fonctions  $e^x$ ,  $x^n$ ,  $\sin x$ , etc. sont infiniment différentiables mais ne sont pas à support compact dans tout intervalle  $]a, b[$  ou même dans  $\mathcal{R}$ . Elles n'appartiennent donc pas à  $\mathcal{D}(\Omega)$ . Inversement, la fonction  $f(x) = |x|$  si  $|x| \leq 1$  et 0 partout ailleurs est à support compact (son support est l'intervalle  $[-1, 1]$ ) mais n'est pas différentiable en  $x = 0$  et  $x = \pm 1$ . Cette fonction n'appartient donc pas à  $\mathcal{D}(\Omega)$ . ◀

► **Exemple 2.3**

Les fonctions de  $\mathcal{D}(\Omega)$  ne sont pas triviales à construire. L'exemple le plus simple en dimension 1 est la fonction :

$$\phi(x) = \begin{cases} \exp\left(\frac{R^2}{|x-p|^2 - R^2}\right) & \text{si } |x-p| < R \\ 0 & \text{si } |x-p| \geq R \end{cases}$$

centrée en  $x = p$  et de rayon  $R$  et que l'on peut voir à la figure 2.1 dans le cas où  $R = 2$  et  $p = 0$ . Le support de cette fonction est l'intervalle fermé  $[p - R, p + R]$ . On choisit les paramètres  $R$  et  $p$  de sorte que ce support soit inclus dans le domaine  $\Omega$ . Sur l'agrandissement de la figure 2.1, on constate aisément la transition très douce (infiniment différentiable) en  $x = 2$  ( $x = R$  dans le cas général). On peut généraliser en plusieurs dimensions cette fonction par la formule suivante :

$$\phi(\mathbf{x}) = \begin{cases} \exp\left(\frac{R^2}{\|\mathbf{x}-\mathbf{p}\|_2^2 - R^2}\right) & \text{si } \|\mathbf{x}-\mathbf{p}\|_2 < R \\ 0 & \text{si } \|\mathbf{x}-\mathbf{p}\|_2 \geq R \end{cases} \quad (2.2)$$

Le point  $\mathbf{p}$  désigne en quelque sorte le centre de la fonction et  $R$  le rayon de la bulle. Ici encore, le disque fermé de centre  $\mathbf{p}$  et de rayon  $R$  doit être inclus dans  $\Omega$ . La figure 2.2 illustre en dimension 2 le cas où  $R = 2$  et  $\mathbf{p}$  est situé à l'origine. On peut ainsi varier le point  $\mathbf{p}$  et le rayon  $R$  pour obtenir toute une famille de fonctions de  $\mathcal{D}(\Omega)$ . ◀

Pour arriver aux distributions, il est nécessaire d'introduire un peu de terminologie. Une fonction  $\psi$  est une application qui associe à un point  $x$  un scalaire  $\psi(x)$ . On peut généraliser cette définition et ainsi introduire la notion de fonctionnelle.

**Définition 2.4**

Une *fonctionnelle*  $T$  est une application qui associe à une fonction  $\psi(x)$  d'un ensemble  $E$ , un scalaire noté  $\langle T, \psi \rangle$ . L'ensemble de fonctions  $E$  est appelé le *domaine de définition de la fonctionnelle*.

Une fonctionnelle est une fonction agissant sur des fonctions c.-à-d. une « fonction de fonctions ». Les exemples abondent de fonctionnelles qui nous seront très utiles. En voici quelques unes.

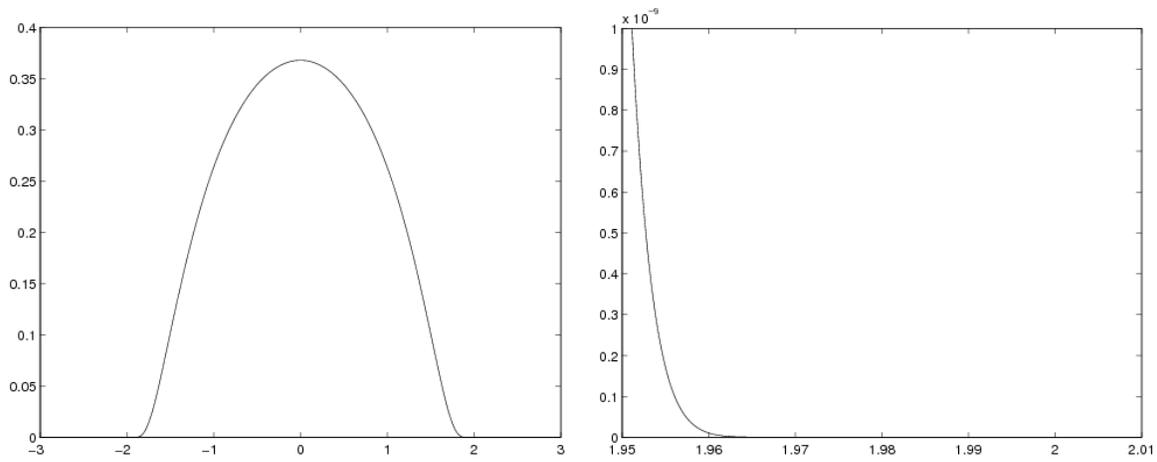


Figure 2.1 – Fonction  $\phi(x) = e^{\frac{R^2}{(x-p)^2 - R^2}}$  ( $R = 2$ ,  $p = 0$ )

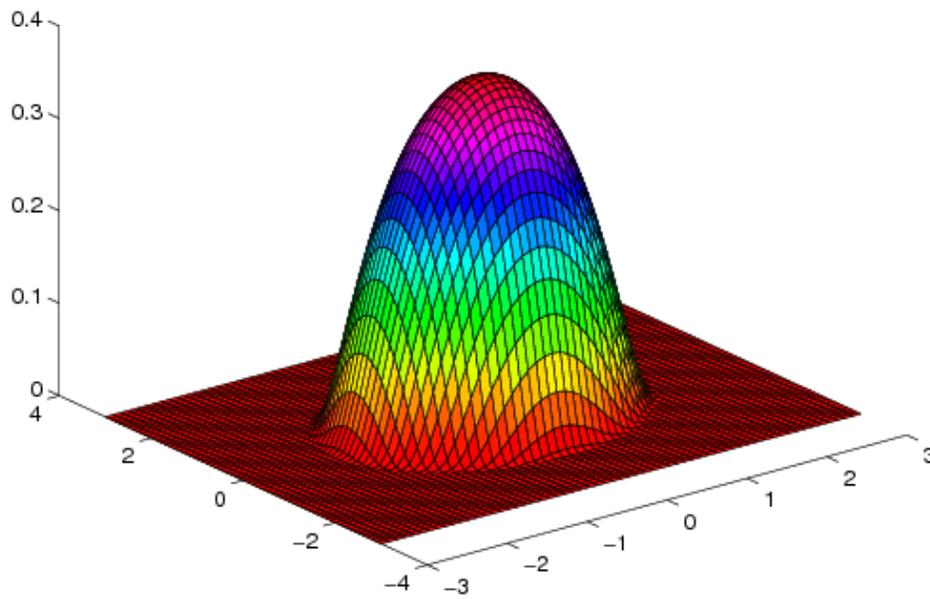


Figure 2.2 – Fonction  $\phi(x) = e^{\frac{R^2}{x_1^2 + x_2^2 - R^2}}$  ( $R = 2$  et  $p = 0$ )

► **Exemple 2.4**

Si  $\psi(\mathbf{x})$  est une fonction intégrable sur  $\Omega$ , on peut alors définir :

$$\langle T_1, \psi \rangle \stackrel{\text{déf.}}{=} \int_{\Omega} \psi(\mathbf{x}) \, dv \quad (2.3)$$

Le domaine de définition de cette première fonctionnelle est donc l'ensemble des fonctions intégrables sur  $\Omega$  c.-à-d. les fonctions pour lesquelles :

$$\int_{\Omega} \psi(\mathbf{x}) \, dv < \infty$$

Notre deuxième exemple est la « fonction » de Dirac définie par :

$$\langle \delta_{\mathbf{a}}, \psi \rangle \stackrel{\text{déf.}}{=} \psi(\mathbf{a}) \quad (2.4)$$

Nous verrons plus loin qu'il ne s'agit pas du tout d'une fonction. Par contre, il s'agit bien d'une fonctionnelle dont le domaine de définition est l'ensemble des fonctions définies au point  $\mathbf{a}$ .

Enfin, si  $f(\mathbf{x})$  est une fonction donnée, on peut alors définir une autre fonctionnelle de la façon suivante :

$$\langle T_f, \psi \rangle \stackrel{\text{déf.}}{=} \int_{\Omega} f(\mathbf{x})\psi(\mathbf{x}) \, dv \quad (2.5)$$

Il nous faudra préciser un peu plus tard le domaine de définition de cette fonctionnelle c.-à-d. pour quelles fonctions  $f(\mathbf{x})$  et  $\psi(\mathbf{x})$  une telle expression a un sens. On remarque que la fonctionnelle  $T_1$  correspond au cas  $f(x) = 1$ . ◀

◆ **Remarque 2.1**

Dans l'exemple précédent, la définition même des intégrales est imprécise. Pour leur donner un sens rigoureux, il faut faire appel à la théorie de la *mesure* et plus particulièrement à l'*intégrale de Lebesgue* (voir par exemple Bartle, réf. [3] ou Jean, réf. [23]). La théorie de Lebesgue donne un sens précis à des expressions de la forme :

$$\int_{\Omega} \psi(\mathbf{x}) \, dv \quad \text{et} \quad \int_{\partial\Omega} \psi(s) \, ds$$

et en particulier à  $dv$  et  $ds$  qui sont alors des mesures. Le lecteur non familier avec cette théorie peut voir l'intégrale de Lebesgue comme une généralisation de l'intégration classique de Riemann. Nous ne ferons dans cet ouvrage que quelques références occasionnelles à la notion de mesure. ◆

► **Exemple 2.5**

On peut également définir des fonctionnelles agissant sur un vecteur de fonctions :

$$\mathbf{w} = (w_1(x_1, x_2, x_3), w_2(x_1, x_2, x_3), w_3(x_1, x_2, x_3))$$

On considère par exemple un corps fait d'un matériau isotrope (en dimension 3) et soumis à des forces externes (forces par unité de volume en  $\text{N/m}^3$ ) :

$$\mathbf{r}(\mathbf{x}) = \mathbf{r}(x_1, x_2, x_3) = (r_1(x_1, x_2, x_3), r_2(x_1, x_2, x_3), r_3(x_1, x_2, x_3))$$

Sous l'effet de ces sollicitations, le corps se déformera de manière à minimiser la *fonctionnelle énergie* définie par :

$$J(\mathbf{w}) = \frac{1}{2} \int_{\Omega} \left( \lambda \left( \sum_{i=1}^3 \frac{\partial w_i}{\partial x_i} \right)^2 + \sum_{i,j=1}^3 \frac{\mu}{4} \left( \frac{\partial w_i}{\partial x_j} + \frac{\partial w_j}{\partial x_i} \right)^2 \right) dv - \int_{\Omega} \mathbf{r} \cdot \mathbf{w} dv \quad (2.6)$$

où  $\lambda$  et  $\mu$  sont les coefficients de Lamé du matériau (en pascals).

Le premier terme de cette expression est l'énergie de déformation tandis que le deuxième terme correspond à l'énergie potentielle des forces externes. On cherchera donc, parmi les déplacements admissibles  $\mathbf{w}$  (en m), celui qui minimise cette fonctionnelle non linéaire. Les unités de  $J(\mathbf{w})$  sont des N m, ce qui correspond bien à une énergie. ◀

#### Définition 2.5

Une fonctionnelle  $T$  est dite *linéaire* sur l'ensemble de fonctions  $E$  si elle vérifie les conditions suivantes :

1.  $\langle T, \psi_1 + \psi_2 \rangle = \langle T, \psi_1 \rangle + \langle T, \psi_2 \rangle, \quad \forall \psi_1 \text{ et } \forall \psi_2 \in E;$
2.  $\langle T, \beta \psi \rangle = \beta \langle T, \psi \rangle, \quad \forall \beta \in \mathbf{R} \text{ et } \forall \psi \in E;$

On peut contracter ces deux conditions en vérifiant que  $\forall \alpha_1, \alpha_2 \in \mathbf{R}$  et  $\forall \psi_1, \psi_2 \in E$ , on a :

$$\langle T, \alpha_1 \psi_1 + \alpha_2 \psi_2 \rangle = \alpha_1 \langle T, \psi_1 \rangle + \alpha_2 \langle T, \psi_2 \rangle$$

Les fonctionnelles 2.3, 2.4 et 2.5 sont linéaires, comme on peut facilement le constater. Prenons par exemple la fonctionnelle  $T_f$ . On a alors :

$$\langle T_f, \beta \psi \rangle = \int_{\Omega} f(\mathbf{x}) \beta \psi(\mathbf{x}) dv = \beta \int_{\Omega} f(\mathbf{x}) \psi(\mathbf{x}) dv = \beta \langle T_f, \psi \rangle$$

De plus :

$$\begin{aligned} \langle T_f, \psi_1 + \psi_2 \rangle &= \int_{\Omega} f(\mathbf{x}) (\psi_1(\mathbf{x}) + \psi_2(\mathbf{x})) dv = \int_{\Omega} f(\mathbf{x}) \psi_1(\mathbf{x}) dv + \int_{\Omega} f(\mathbf{x}) \psi_2(\mathbf{x}) dv \\ &= \langle T_f, \psi_1 \rangle + \langle T_f, \psi_2 \rangle \end{aligned}$$

Par contre, la fonctionnelle :

$$\langle T, \psi \rangle = \int_{\Omega} (\psi(\mathbf{x}))^2 dv$$

ne l'est pas puisque  $(\psi_1 + \psi_2)^2 \neq \psi_1^2 + \psi_2^2$ . Il en est de même pour la fonctionnelle 2.6 qui est en fait quadratique comme nous le verrons plus loin.

◆ **Remarque 2.2**

Pour compléter le tableau, il nous faudrait introduire la notion de *continuité* d'une fonctionnelle. Pour cela, il serait nécessaire d'introduire une *topologie* sur l'ensemble  $E$  et de définir la notion de *fonctionnelle continue*. Cela ne nous paraît pas essentiel à ce stade-ci pour bien comprendre la suite. Nous reviendrons un peu plus loin sur la notion de continuité dans le cadre des espaces de Hilbert. ◆

**Définition 2.6**

Une fonctionnelle linéaire (et continue) définie sur l'espace  $\mathcal{D}(\Omega)$  est une *distribution*. L'ensemble des distributions possède une structure d'espace vectoriel que nous notons  $\mathcal{D}'(\Omega)$ .

L'espace  $\mathcal{D}'(\Omega)$  est aussi appelé *espace dual* de  $\mathcal{D}(\Omega)$ . Nous reviendrons sur la notion d'espace dual un peu plus loin. Les trois fonctionnelles linéaires que nous avons vues sont des distributions. En particulier, la *distribution de Dirac* :

$$\langle \delta_{\mathbf{a}}, \phi \rangle = \phi(\mathbf{a}) \quad \forall \phi \in \mathcal{D}(\Omega) \quad (2.7)$$

est sans doute la plus connue.

◆ **Remarque 2.3**

Bien que cela soit souvent le cas, les distributions ne sont pas toujours associées à ce qu'on pourrait qualifier de forme intégrale comme cela est le cas pour les distributions  $T_1$  et  $T_f$ . Ainsi, la distribution de Dirac ne peut s'écrire sous forme intégrale bien que l'on trouve parfois la notation :

$$\int_{\Omega} \delta_{\mathbf{a}} \phi dv = \phi(\mathbf{a})$$

Nous verrons un peu plus loin qu'il n'existe pas de fonction « classique » vérifiant une telle propriété. Cette notation est donc erronée et ne devrait pas être employée. *Il faut donc éviter d'associer automatiquement l'expression  $\langle T, \phi \rangle$  avec une forme intégrale.* ◆

On peut manipuler les distributions, un peu comme les fonctions et définir l'addition, la multiplication d'une distribution par un scalaire et l'égalité de deux distributions.

– **Addition de deux distributions :**

On définit l'addition de deux distributions  $T_1$  et  $T_2$  par :

$$\langle T_1 + T_2, \phi \rangle \stackrel{\text{déf.}}{=} \langle T_1, \phi \rangle + \langle T_2, \phi \rangle \quad \forall \phi \in \mathcal{D}(\Omega) \quad (2.8)$$

– **Multiplication d'une distribution  $T$  par un scalaire  $\beta \in R$  :**

$$\langle \beta T, \phi \rangle \stackrel{\text{déf.}}{=} \beta \langle T, \phi \rangle \quad \forall \phi \in \mathcal{D}(\Omega) \quad (2.9)$$

– **Multiplication d'une distribution  $T$  par une fonction infiniment différentiable  $h(\mathbf{x})$  :**

$$\langle hT, \phi \rangle \stackrel{\text{déf.}}{=} \langle T, h\phi \rangle \quad \forall \phi \in \mathcal{D}(\Omega) \quad (2.10)$$

– **Égalité de deux distributions :**

Les distributions  $T_1$  et  $T_2$  sont dites égales si :

$$\langle T_1, \phi \rangle = \langle T_2, \phi \rangle \quad \forall \phi \in \mathcal{D}(\Omega) \quad (2.11)$$

On remarque que les distributions ne sont définies que par l'effet qu'elles produisent sur les fonctions de  $\mathcal{D}(\Omega)$ . Les définitions paraissent alors naturelles et il est important de noter que deux distributions sont réputées égales si elles ont le même effet sur toutes les fonctions de  $\mathcal{D}(\Omega)$ .

## Fonctions localement intégrables et distributions régulières

### Définition 2.7

Une fonction  $f(\mathbf{x})$  est dite *localement intégrable* si elle est intégrable sur tout compact inclus dans  $\Omega$  ou encore si :

$$\int_A |f(\mathbf{x})| dv < \infty \quad \forall \text{ compact } A \subset \Omega \quad (2.12)$$

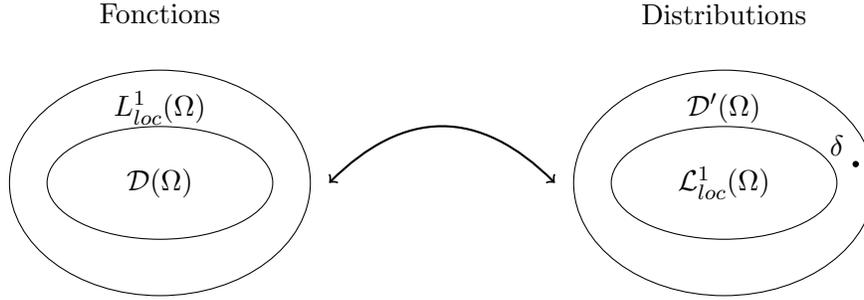
L'ensemble des fonctions localement intégrables forme un espace noté  $L^1_{loc}(\Omega)$ .

Notons immédiatement qu'une fonction peut ne pas être intégrable sur tout le domaine  $\Omega$  mais être tout de même localement intégrable. On pense par exemple à la fonction  $f(x) = 1$  lorsque  $\Omega = R$ .

On peut associer à une fonction localement intégrable  $f(\mathbf{x}) \in L^1_{loc}(\Omega)$  une distribution  $T_f$  définie par :

$$\langle T_f, \phi \rangle = \int_{\Omega} f(\mathbf{x})\phi(\mathbf{x}) dv \quad \forall \phi \in \mathcal{D}(\Omega)$$

Il est facile de constater que l'expression précédente a un sens c.-à-d. :

Figure 2.3 – Espaces de fonctions  $L^1_{loc}(\Omega)$  et de distributions  $\mathcal{L}^1_{loc}(\Omega)$ 

$$\left| \int_{\Omega} f(\mathbf{x})\phi(\mathbf{x}) dv \right| \leq \int_{\Omega} |f(\mathbf{x})\phi(\mathbf{x})| dv \leq M_{\phi} \int_{K_{\phi}} |f(\mathbf{x})| dv < \infty$$

car l'intégrale porte en fait sur le support compact  $K_{\phi}$  de la fonction  $\phi$  qui y est bornée par  $M_{\phi}$ . Ainsi, à toute fonction localement intégrable, nous avons associé une distribution. Une telle distribution est dite *régulière*. Nous noterons  $\mathcal{L}^1_{loc}(\Omega)$  l'espace des distributions associées à une fonction de  $L^1_{loc}(\Omega)$ . Inversement, pour une distribution quelconque, il n'est pas toujours possible de lui associer une fonction localement intégrable. L'espace  $\mathcal{L}^1_{loc}(\Omega)$  est donc un sous-espace propre de  $\mathcal{D}'(\Omega)$  tel qu'illustré à la figure 2.3.

### ♣ Lemme 2.1

La distribution de Dirac n'est pas régulière.

#### Démonstration :

On doit montrer qu'il n'existe pas de fonction  $f(\mathbf{x})$  localement intégrable telle que :

$$\langle \delta_{\mathbf{a}}, \phi \rangle = \int_{\Omega} f(\mathbf{x})\phi(\mathbf{x}) dv = \phi(\mathbf{a}) \quad \forall \phi \in \mathcal{D}(\Omega)$$

Si une telle fonction  $f(\mathbf{x})$  existait, en posant  $\phi(\mathbf{x}) = \phi_n(\mathbf{x})$ , où  $\phi_n(\mathbf{x})$  est la fonction définie par l'équation 2.2 centrée en  $\mathbf{a}$  ( $\mathbf{p} = \mathbf{a}$ ) et de rayon  $1/n$  ( $R = 1/n$ ), on aurait :

$$\int_{\Omega} f(\mathbf{x})\phi_n(\mathbf{x}) dv = \phi_n(\mathbf{a}) = e^{-1}$$

Si  $C(\mathbf{a}, n)$  désigne le disque de centre  $\mathbf{a}$  et de rayon  $1/n$ , on aurait de plus :

$$e^{-1} = |\phi_n(\mathbf{a})| \leq \int_{C(\mathbf{a}, n)} |f(\mathbf{x})| |\phi_n(\mathbf{x})| dv \leq e^{-1} \int_{C(\mathbf{a}, n)} |f(\mathbf{x})| dv$$

puisque les fonctions  $\phi_n(\mathbf{x})$  sont toutes bornées par  $e^{-1}$ . La dernière inégalité constitue une contradiction puisque le terme de droite tend vers 0 lorsque  $n$  tend vers l'infini. ♣

## Dérivées d'une distribution

On peut généraliser la notion de dérivée aux distributions, ce qui nous sera très utile pour les équations aux dérivées partielles. Nous nous limiterons pour l'instant aux fonctions d'une variable et nous supposons que l'ouvert  $\Omega$  est tout simplement l'intervalle ouvert  $]a, b[$ . Soit donc une fonction continue  $f(x)$  dont la dérivée est *continue par morceaux* c.-à-d. dont la dérivée possède éventuellement un nombre fini de *discontinuités de première espèce* (sauts de hauteur finie) dans l'intervalle  $]a, b[$ . On associe alors à la dérivée  $f'(x)$  une distribution  $T_{f'}$  définie par :

$$\langle T_{f'}, \phi \rangle = \int_a^b f'(x)\phi(x)dx \quad \forall \phi \in \mathcal{D}(]a, b[)$$

On peut alors intégrer par parties et écrire :

$$\begin{aligned} \langle T_{f'}, \phi \rangle &= - \int_a^b f(x)\phi'(x)dx + f(b)\phi(b) - f(a)\phi(a) \\ &= - \int_a^b f(x)\phi'(x)dx \end{aligned}$$

Les termes de bord de l'intégration par parties s'annulent car les fonctions  $\phi$  sont à support compact dans l'intervalle  $]a, b[$  et doivent donc s'annuler en  $a$  et  $b$ . Il vient alors que :

$$\langle T_{f'}, \phi \rangle = - \int_a^b f(x)\phi'(x)dx = - \langle T_f, \phi' \rangle$$

Le raisonnement précédent n'est valide que pour les fonctions  $f(x)$  dont la dérivée est continue par morceaux. Par contre, nous nous servons de la dernière relation comme définition de la dérivée d'une distribution. L'idée derrière cela est que ce qui est vrai pour les fonctions usuelles doit être encore valide pour les distributions. On peut faire un raisonnement similaire, sous des hypothèses de régularité appropriées sur la fonction  $f(x)$ , pour traiter les dérivées d'ordre supérieur. Il en résulte alors de manière naturelle la définition suivante.

### Définition 2.8

On définit la dérivée d'une distribution  $T$  par la relation :

$$\langle T', \phi \rangle \stackrel{\text{déf.}}{=} - \langle T, \phi' \rangle \quad \forall \phi \in \mathcal{D}(]a, b[) \quad (2.13)$$

On définit de manière similaire, la dérivée d'ordre  $n$  :

$$\langle T^{(n)}, \phi \rangle \stackrel{\text{déf.}}{=} (-1)^n \langle T, \phi^{(n)} \rangle \quad \forall \phi \in \mathcal{D}(]a, b[) \quad (2.14)$$

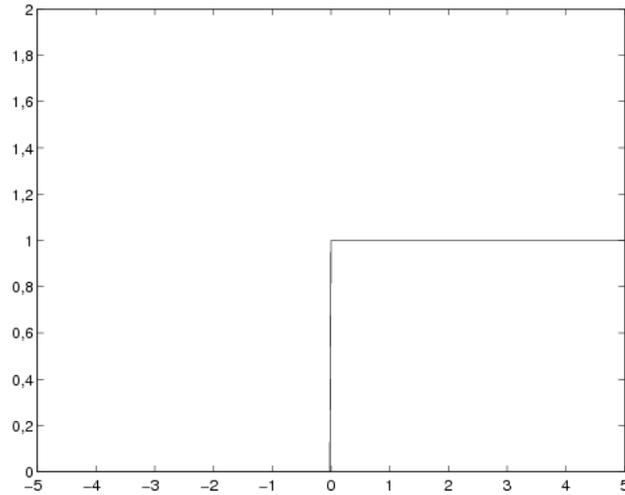


Figure 2.4 – Fonction de Heaviside

La dérivée de la distribution  $T$  est passée, via l'intégration par parties aux fonctions  $\phi$  qui sont infiniment différentiables. Il en résulte que, contrairement aux fonctions, les distributions sont toutes infiniment différentiables.

► **Exemple 2.6**

Nous supposons dans cet exemple que l'ouvert  $\Omega$  est  $] -\infty, \infty[$ . La fonction de Heaviside ou fonction échelon s'écrit :

$$H(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x > 0 \end{cases}$$

et est illustrée à la figure 2.4. La fonction n'est pas définie en 0 et n'est bien sûr pas dérivable au sens classique à cet endroit. Cette fonction étant localement intégrable, on lui associe une distribution que nous notons  $T_H$  définie par :

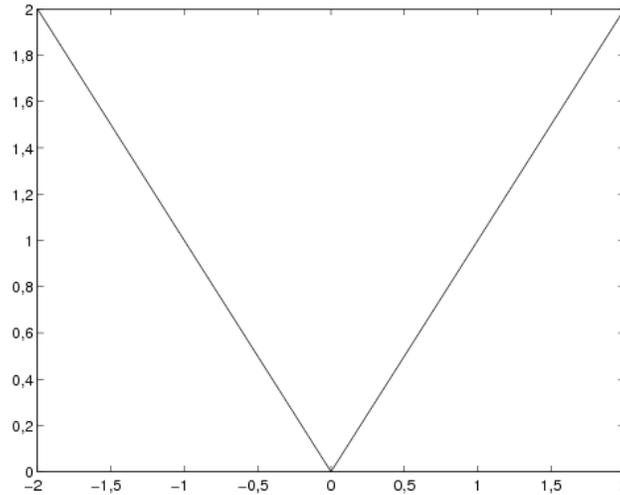
$$\langle T_H, \phi \rangle = \int_{-\infty}^{\infty} H(x)\phi(x)dx = \int_0^{\infty} \phi(x)dx$$

et dont la dérivée est à son tour définie à l'aide de la relation 2.13 :

$$\begin{aligned} \langle T'_H, \phi \rangle &\stackrel{\text{déf.}}{=} - \langle T_H, \phi'(x) \rangle = - \int_0^{\infty} \phi'(x)dx \\ &= -(\phi(\infty) - \phi(0)) = \phi(0) \quad \forall \phi \in \mathcal{D}(\Omega) \end{aligned}$$

On en conclut immédiatement que :

$$\langle T'_H, \phi \rangle = \phi(0) = \langle \delta_0, \phi \rangle \quad \forall \phi \in \mathcal{D}(\Omega)$$

Figure 2.5 – Fonction  $f(x) = |x|$ 

et donc, en vertu de la relation 2.11, la dérivée au sens des distributions de la fonction de Heaviside est tout simplement la distribution de Dirac en 0. ◀

► **Exemple 2.7**

Les distributions étant toutes différentiables, on s'intéresse à la dérivée de la distribution de Dirac. On a ainsi :

$$\langle \delta'_{\mathbf{a}}, \phi \rangle \stackrel{\text{déf.}}{=} - \langle \delta_{\mathbf{a}}, \phi' \rangle = -\phi'(\mathbf{a}) \quad \forall \phi \in \mathcal{D}(\Omega)$$

◀

◆ **Remarque 2.4**

Il est tentant d'identifier une fonction localement intégrable et la distribution régulière qui lui est associée. Ainsi la fonction de Heaviside serait confondue avec la distribution  $T_H$  définie plus haut. Nous essaierons d'éviter cette association qui peut mener à des abus. ◆

► **Exemple 2.8**

La fonction  $f(x) = |x|$  n'est pas différentiable en 0, comme l'illustre la figure 2.5. On peut par contre lui associer une distribution  $T_{|x|}$  définie par :

$$\langle T_{|x|}, \phi \rangle = \int_{-\infty}^{\infty} |x| \phi(x) dx$$

qui elle est infiniment différentiable et dont la dérivée première est aussi définie à l'aide de la relation 2.13. Au sens des distributions, on a :

$$\begin{aligned}
\langle T'_{|x|}, \phi \rangle &\stackrel{\text{déf.}}{=} - \langle T_{|x|}, \phi' \rangle = - \int_{-\infty}^{\infty} |x| \phi'(x) dx \\
&= - \left( \int_{-\infty}^0 -x \phi'(x) dx + \int_0^{\infty} x \phi'(x) dx \right) \\
&= x \phi(x) \Big|_{-\infty}^0 - \int_{-\infty}^0 \phi(x) dx - x \phi(x) \Big|_0^{\infty} + \int_0^{\infty} \phi(x) dx \\
&= - \int_{-\infty}^0 \phi(x) dx + \int_0^{\infty} \phi(x) dx = \int_{-\infty}^{\infty} g(x) \phi(x) dx \\
&= \langle T_g, \phi \rangle \quad \forall \phi \in \mathcal{D}(\Omega)
\end{aligned}$$

où  $T_g$  est la distribution régulière associée à la fonction localement intégrable :

$$g(x) = \begin{cases} -1 & \text{si } x < 0 \\ +1 & \text{si } x > 0 \end{cases}$$

et en ce sens,  $T'_{|x|} = T_g$ . Très souvent, on note abusivement  $|x|' = g$  ce qui n'est pas rigoureusement vrai. Regardons maintenant la dérivée seconde  $T''_{|x|}$ . Au sens des distributions et de l'équation 2.14 avec  $n = 2$ , on a :

$$\begin{aligned}
\langle T''_{|x|}, \phi \rangle &\stackrel{\text{déf.}}{=} (-1)^2 \langle T_{|x|}, \phi'' \rangle = \int_{-\infty}^{\infty} |x| \phi''(x) dx \\
&= \left( \int_{-\infty}^0 -x \phi''(x) dx + \int_0^{\infty} x \phi''(x) dx \right) \\
&= -x \phi'(x) \Big|_{-\infty}^0 + \int_{-\infty}^0 \phi'(x) dx + x \phi'(x) \Big|_0^{\infty} - \int_0^{\infty} \phi'(x) dx \\
&= \int_{-\infty}^0 \phi'(x) dx - \int_0^{\infty} \phi'(x) dx \\
&= (\phi(0) - \phi(-\infty)) - (\phi(\infty) - \phi(0)) = 2\phi(0) \\
&= \langle 2\delta_0, \phi \rangle \quad \forall \phi \in \mathcal{D}(\Omega)
\end{aligned}$$

La dérivée seconde de la distribution associée à la fonction valeur absolue est donc tout simplement deux fois la distribution de Dirac en 0. Ici encore, on note parfois abusivement  $|x|'' = 2\delta_0$ , notation qu'il convient d'éviter le plus possible. ◀

On peut maintenant passer à un cas général que nous ne démontrons pas mais qui est facile à déduire de la discussion et des exemples précédents.

★ **Théorème 2.1**

Soit  $f(x)$  une fonction continue et dérivable sauf aux points  $x_i, i = 1, 2, \dots, n$  où elle possède des discontinuités de première espèce (c.-à-d. les limites à gauche et à droite de  $f(x)$  existent et sont finies). La dérivée au sens des distributions de  $f(x)$  s'écrit :

$$T_f' = T_{f'} + \sum_{i=1}^n s_i^f \delta_{x_i} \quad (2.15)$$

où  $s_i^f$  est le saut de la fonction  $f(x)$  en  $x = x_i$  qui est défini par :

$$s_i^f = f(x_i^+) - f(x_i^-)$$

★

◆ **Remarque 2.5**

Il faut savoir interpréter correctement ce résultat. Il nous assure que la dérivée de la distribution associée à  $f(x)$  n'est autre que la distribution associée à la fonction  $f'(x)$  (partout où cette dérivée existe) plus une contribution provenant des sauts de  $f(x)$  et faisant intervenir les distributions de Dirac aux points correspondants. De toute évidence, si la fonction  $f(x)$  est continue, alors  $T_f' = T_{f'}$ .

La dérivée de la fonction de Heaviside entre dans le cadre de ce théorème. Cette fonction possède un saut de hauteur 1 en  $x = 0$  et de plus,  $T_{H'} = 0$ , ce qui signifie que  $T_H' = \delta_0$ . ◆

### 2.1.2 Distributions de plusieurs variables

Nous pouvons maintenant revenir au cas plus général en dimension  $n$ . Soit  $\mathbf{m} = (m_1, m_2, \dots, m_n)$  un multi-indice. La dérivée partielle d'une fonction de plusieurs variables  $f(x_1, x_2, \dots, x_n)$  s'écrit :

$$\frac{\partial^{\mathbf{m}} f}{\partial \mathbf{x}^{\mathbf{m}}} = \frac{\partial^{|\mathbf{m}|} f}{\partial x_1^{m_1} \partial x_2^{m_2} \dots \partial x_n^{m_n}}$$

où  $|\mathbf{m}| = m_1 + m_2 + \dots + m_n$ .

Suivant cette notation, on peut définir les dérivées partielles d'une distribution.

**Définition 2.9**

La dérivée partielle de la distribution  $T$  par rapport au multi-indice  $\mathbf{m}$  est :

$$\left\langle \frac{\partial^{\mathbf{m}} T}{\partial \mathbf{x}^{\mathbf{m}}}, \phi \right\rangle \stackrel{\text{déf.}}{=} (-1)^{|\mathbf{m}|} \left\langle T, \frac{\partial^{\mathbf{m}} \phi}{\partial \mathbf{x}^{\mathbf{m}}} \right\rangle \quad \forall \phi \in \mathcal{D}(\Omega) \quad (2.16)$$

Terminons cette section par quelques exemples dans le cas bidimensionnel.

**► Exemple 2.9**

La *simple couche* est en quelque sorte la généralisation en 2 dimensions de la distribution de Dirac. Soit  $C$  une courbe du plan et  $f(\mathbf{x})$  une fonction intégrable sur  $C$ . On pose alors :

$$\left\langle \delta_C^f, \phi \right\rangle \stackrel{\text{déf.}}{=} \int_C f \phi \, ds \quad (2.17)$$

Il est facile de vérifier que  $\delta_C^f$  est une distribution. En fait, la généralisation de la distribution de Dirac correspond au cas où  $f(\mathbf{x}) = 1$ . ◀

**► Exemple 2.10**

On peut sans difficulté parler de distribution vectorielle de la forme  $\mathbf{T} = (T_1, T_2, T_3)$  où chaque composante  $T_i$  est une distribution. On a alors :

$$\left\langle \mathbf{T}, \phi \right\rangle = \sum_{i=1}^3 \left\langle T_i, \phi_i \right\rangle \quad \forall \phi = (\phi_1, \phi_2, \phi_3) \in (\mathcal{D}(\Omega))^3$$

On peut alors généraliser certains résultats connus pour les fonctions vectorielles. ◀

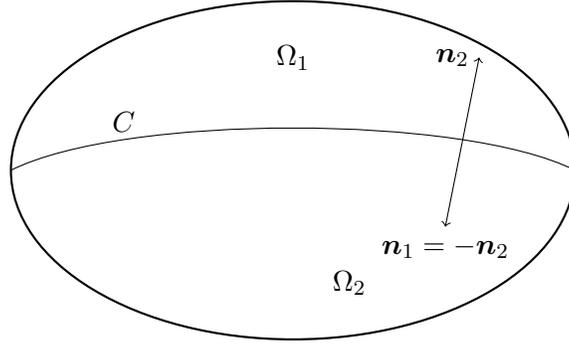
**► Exemple 2.11**

En plusieurs dimensions, les opérateurs essentiels de dérivation sont le gradient, la divergence et le laplacien. Dans le cas des distributions, par analogie avec la relation A.3 valide pour les fonctions suffisamment régulières, on définit la divergence d'une distribution par l'équation :

$$\left\langle \nabla \cdot \mathbf{T}, \phi \right\rangle \stackrel{\text{déf.}}{=} - \left\langle \mathbf{T}, \nabla \phi \right\rangle \quad \forall \phi \in \mathcal{D}(\Omega) \quad (2.18)$$

en rappelant que les fonctions  $\phi$  s'annulent sur la frontière de  $\Omega$ . ◀

**► Exemple 2.12**

Figure 2.6 – Domaines  $\Omega_1$  et  $\Omega_2$  séparés par une courbe  $C$ 

Soit un champ de vecteurs  $\mathbf{u}(\mathbf{x})$  défini sur un ouvert  $\Omega = \Omega_1 \cup \Omega_2 \subset \mathcal{R}^3$  par la relation :

$$\mathbf{u}(\mathbf{x}) = \begin{cases} \mathbf{u}_1(\mathbf{x}) & \text{si } \mathbf{x} \in \Omega_1 \\ \mathbf{u}_2(\mathbf{x}) & \text{si } \mathbf{x} \in \Omega_2 \end{cases}$$

On suppose les deux sous-domaines  $\Omega_1$  et  $\Omega_2$  de frontière  $\Gamma_1$  et  $\Gamma_2$  séparés par une surface  $C$  (ou une courbe en dimension 2 tel qu'illustré à la figure 2.6). On suppose de plus que le champ  $\mathbf{u}$  est dérivable dans le complémentaire de  $C$  (c.-à-d. partout sauf sur la courbe  $C$ ) et que sur cette surface, les champs  $\mathbf{u}_1(\mathbf{x})$  et  $\mathbf{u}_2(\mathbf{x})$  sont discontinus et possèdent un saut de hauteur  $\mathbf{s} = \mathbf{u}_2 - \mathbf{u}_1$ .

On peut bien sûr associer à  $\mathbf{u}$  une distribution régulière que nous noterons  $T_{\mathbf{u}}$ . On a alors en vertu de la relation 2.18 :

$$\begin{aligned} \langle \nabla \cdot T_{\mathbf{u}}, \phi \rangle &\stackrel{\text{déf.}}{=} - \langle T_{\mathbf{u}}, \nabla \phi \rangle = - \int_{\Omega} \mathbf{u} \cdot \nabla \phi \, dv \\ &= - \int_{\Omega_1} \mathbf{u}_1 \cdot \nabla \phi \, dv - \int_{\Omega_2} \mathbf{u}_2 \cdot \nabla \phi \, dv \end{aligned}$$

On peut appliquer la relation A.3 à chacune des deux dernières intégrales pour obtenir :

$$\begin{aligned} \langle \nabla \cdot T_{\mathbf{u}}, \phi \rangle &= \int_{\Omega_1} \nabla \cdot \mathbf{u}_1 \phi \, dv - \int_{\Gamma_1} \mathbf{u}_1 \cdot \mathbf{n}_1 \phi \, ds \\ &\quad + \int_{\Omega_2} \nabla \cdot \mathbf{u}_2 \phi \, dv - \int_{\Gamma_2} \mathbf{u}_2 \cdot \mathbf{n}_2 \phi \, ds \end{aligned}$$

Les fonctions  $\phi$  s'annulent identiquement sur la frontière de  $\Omega$  et les intégrales surfaciques ne portent

en définitive que sur la surface  $C$ . On a alors :

$$\begin{aligned} \langle \nabla \cdot T\mathbf{u}, \phi \rangle &= \int_{\Omega_1} \nabla \cdot \mathbf{u}_1 \phi \, dv - \int_C \mathbf{u}_1 \cdot \mathbf{n}_1 \phi \, ds \\ &+ \int_{\Omega_2} \nabla \cdot \mathbf{u}_2 \phi \, dv - \int_C \mathbf{u}_2 \cdot \mathbf{n}_2 \phi \, ds \end{aligned}$$

On choisit maintenant de manière arbitraire de poser sur la surface  $C$ ,  $\mathbf{n} = \mathbf{n}_1 = -\mathbf{n}_2$  et on trouve :

$$\langle \nabla \cdot T\mathbf{u}, \phi \rangle = \int_{\Omega_1} \nabla \cdot \mathbf{u}_1 \phi \, dv + \int_{\Omega_2} \nabla \cdot \mathbf{u}_2 \phi \, dv + \int_C (\mathbf{s} \cdot \mathbf{n}) \phi \, ds$$

Il en résulte le résultat très important suivant :

$$\begin{aligned} \langle \nabla \cdot T\mathbf{u}, \phi \rangle &= \sum_{i=1}^2 \int_{\Omega_i} \nabla \cdot \mathbf{u}_i \phi \, dv + \langle \delta_C^{\mathbf{s} \cdot \mathbf{n}}, \phi \rangle \\ &= \langle T_{\nabla} \cdot \mathbf{u} + \delta_C^{\mathbf{s} \cdot \mathbf{n}}, \phi \rangle \end{aligned} \tag{2.19}$$

◀

L'interprétation de ce résultat est importante. Si un champ de vecteurs est discontinu le long d'une interface  $C$ , la divergence de la distribution associée à ce champ comporte deux parties : une première qui n'est rien d'autre que la distribution régulière associée à la divergence de  $\mathbf{u}$  dans chaque sous-domaine (la divergence y étant bien définie), et une autre partie qui tient compte de la discontinuité sur  $C$  par l'entremise d'une simple couche. Il est clair que ce résultat généralise l'équation 2.15.

#### ◆ Remarque 2.6

Notons que le signe du saut du champ  $\mathbf{u}$  à travers l'interface  $C$  est fixé par le choix de la normale  $\mathbf{n}$  à cette même interface et vice versa. Si on change l'orientation de la normale, on change le signe du saut mais le produit  $\mathbf{s} \cdot \mathbf{n}$  ne change pas puisque :

$$(\mathbf{u}_1 - \mathbf{u}_2) \cdot \mathbf{n}_2 = (\mathbf{u}_2 - \mathbf{u}_1) \cdot \mathbf{n}_1$$

◆

## 2.2 Espaces fonctionnels linéaires

### Définition 2.10

Un *espace fonctionnel linéaire* est un ensemble  $S$  de fonctions définies sur un ouvert  $\Omega$  et vérifiant :

1. Si  $u \in S$  et  $\beta \in R$ , alors  $\beta u \in S$  ;
2. Si  $u \in S$  et  $w \in S$ , alors  $(u + w) \in S$  ;

C'est donc un *espace vectoriel de fonctions*.

### ► Exemple 2.13

L'ensemble des fonctions continues sur un ouvert  $\Omega$ , noté  $C^0(\Omega)$ , est un espace fonctionnel linéaire. De même, l'ensemble des fonctions  $k$  fois différentiables, noté  $C^k(\Omega)$ , est aussi un espace fonctionnel linéaire. Par contre, l'ensemble :

$$S = \{u \in C^0(\Omega) | u(0) = 1\}$$

n'est pas un espace fonctionnel linéaire. En effet, si  $u$  et  $w$  sont dans  $S$  alors leur somme vérifie  $(u + w)(0) = 2$  et la somme n'appartient donc pas à  $S$ . ◀

L'un des espaces fonctionnels qui nous sera le plus utile est l'espace  $L^2(\Omega)$ .

### Définition 2.11

On note  $L^2(\Omega)$  l'ensemble des fonctions de carré sommable c.-à-d.

$$L^2(\Omega) = \left\{ u : \Omega \rightarrow R \mid \int_{\Omega} (u(\mathbf{x}))^2 dv < \infty \right\}$$

### ◆ Remarque 2.7

La définition précédente est beaucoup plus subtile qu'elle n'y paraît. La théorie de Lebesgue ne fait pas de distinction entre des fonctions qui ne diffèrent que sur un ensemble dit de *mesure nulle*. Pour fixer les idées, mentionnons simplement qu'un ensemble de mesure nulle contient très peu de points. Les ensembles finis ou dénombrables de points sont de mesure nulle. Par exemple, la fonction :

$$f(x) = \begin{cases} 0 & \text{si } x \text{ est un nombre irrationnel} \\ 1 & \text{si } x \text{ est un nombre rationnel} \end{cases}$$

sera identifiée à la fonction  $f(x) = 0$  car les nombres rationnels sont dénombrables et donc de mesure de Lebesgue nulle.

Par la suite, nous ne ferons pas de distinction entre 2 fonctions qui ne diffèrent que sur un ensemble de mesure nulle. On écrira par exemple :

$$u_1 \stackrel{\text{p.p.}}{=} u_2$$

signifiant ainsi que les fonctions  $u_1$  et  $u_2$  sont égales presque partout (p.p.) c.-à-d. sauf peut-être sur un ensemble de mesure nulle. ♦

### ★ Théorème 2.2

$L^2(\Omega)$  est un espace fonctionnel linéaire. ★

#### Démonstration :

Pour démontrer que  $L^2(\Omega)$  est un espace fonctionnel, il nous faut établir que si  $u$  est de carré sommable, alors  $\beta u$  est de carré sommable. Mais :

$$\int_{\Omega} (\beta u(\mathbf{x}))^2 dv = \beta^2 \int_{\Omega} (u(\mathbf{x}))^2 dv < \infty$$

et la première propriété est démontrée. La deuxième propriété concernant la somme de 2 fonctions requiert un résultat préliminaire qui lui-même est très important.

### ♣ Lemme 2.2 (Inégalité de Cauchy-Schwartz)

Si  $u$  et  $w$  sont des fonctions de  $L^2(\Omega)$ , alors :

$$\left| \int_{\Omega} uw dv \right| \leq \left( \int_{\Omega} u^2 dv \right)^{1/2} \left( \int_{\Omega} w^2 dv \right)^{1/2} \quad (2.20)$$

#### Démonstration :

Soit  $t$  un nombre réel quelconque. La fonction  $(u + tw)^2$  étant positive quel que soit  $t$ , on a :

$$0 \leq \int_{\Omega} (u + tw)^2 dv \quad \forall t \in \mathbb{R}$$

ce qui peut également s'écrire :

$$0 \leq \int_{\Omega} u^2 dv + 2t \int_{\Omega} uw dv + t^2 \int_{\Omega} w^2 dv \quad \forall t \in \mathbb{R}$$

ou encore :

$$0 \leq At^2 + Bt + C \quad \forall t \in \mathbb{R}$$

où :

$$A = \int_{\Omega} w^2 dv, \quad B = 2 \int_{\Omega} uw dv, \quad C = \int_{\Omega} u^2 dv$$

Il s'agit donc d'un polynôme de degré 2 en la variable  $t$  qui est toujours positif. Cela n'est possible que si le discriminant  $B^2 - 4AC$  est négatif ou nul c.-à-d.  $B^2 - 4AC \leq 0$ . En remplaçant les valeurs de  $A$ ,  $B$  et  $C$ , on trouve immédiatement l'inégalité de Cauchy-Schwartz ce qui termine la démonstration du lemme. ♣

Nous sommes en mesure de démontrer que la somme de deux fonctions de  $L^2(\Omega)$  est également dans  $L^2(\Omega)$ . En effet :

$$\begin{aligned} \int_{\Omega} (u+w)^2 dv &= \int_{\Omega} u^2 dv + 2 \int_{\Omega} uw dv + \int_{\Omega} w^2 dv \\ &\leq \int_{\Omega} u^2 dv + 2 \left( \int_{\Omega} u^2 dv \right)^{1/2} \left( \int_{\Omega} w^2 dv \right)^{1/2} + \int_{\Omega} w^2 dv \\ &= \left[ \left( \int_{\Omega} u^2 dv \right)^{1/2} + \left( \int_{\Omega} w^2 dv \right)^{1/2} \right]^2 < \infty \end{aligned}$$

ce qui complète la démonstration du théorème.

### ♠ Corollaire 2.1

On a l'inclusion  $L^2(\Omega) \subset L^1_{loc}(\Omega)$ .

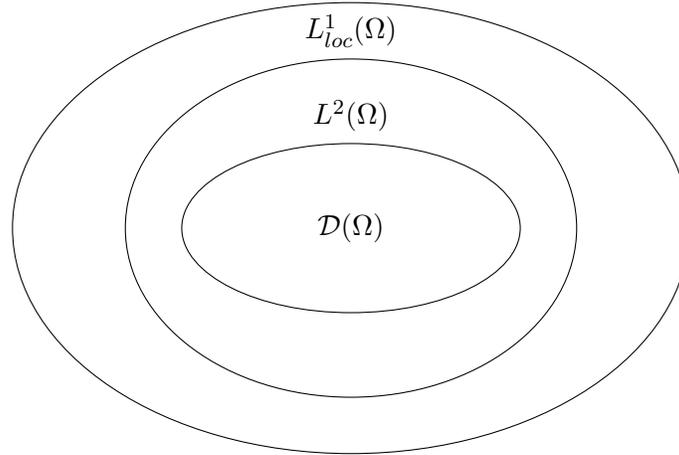
#### Démonstration :

Ce résultat découle de l'inégalité de Cauchy-Schwartz. En effet, si  $f(\mathbf{x}) \in L^2(\Omega)$  et si  $A$  est un compact inclus dans  $\Omega$ , alors :

$$\begin{aligned} \int_A |f(\mathbf{x})| dv &= \int_A |f(\mathbf{x})| 1 dv \leq \left( \int_A f^2(\mathbf{x}) dv \right)^{1/2} \left( \int_A 1^2 dv \right)^{1/2} \\ &= (\text{vol } A)^{1/2} \left( \int_A f^2(\mathbf{x}) dv \right)^{1/2} < \infty \end{aligned}$$

où  $\text{vol } A$  est le volume de  $A$  (ou l'aire suivant la dimension) qui est forcément fini. Les fonctions de  $L^2(\Omega)$  sont en quelque sorte encore plus régulières que celles de  $L^1_{loc}(\Omega)$  comme en témoigne la figure 2.7. ♠

### ► Exemple 2.14

Figure 2.7 – Espace  $L^2(\Omega)$ 

Choisissons, pour les exemples qui suivent,  $\Omega = ]0, 1[$ . La fonction  $f(x) = x^{-1/4}$  est dans  $L^2(]0, 1[)$ , et ce même si elle n'est pas continue en  $x = 0$ . En effet :

$$\int_0^1 (x^{-1/4})^2 dx = \int_0^1 x^{-1/2} dx = 2 < \infty$$

Par contre, la fonction  $f(x) = x^{-1/2}$  n'appartient pas à  $L^2(]0, 1[)$  car :

$$\int_0^1 (x^{-1/2})^2 dx = \int_0^1 x^{-1} dx = \ln x \Big|_0^1 = \infty$$

Toutefois, si on choisit un intervalle  $]a, b[$  ne comprenant pas l'origine, la fonction  $f(x) = x^{-1/2}$  sera dans  $L^2(]a, b[)$ . ◀

La notion d'espace fonctionnel linéaire n'est pas suffisante pour atteindre notre objectif de résoudre les équations aux dérivées partielles. Il nous faut introduire une métrique sur les espaces fonctionnels qui nous permettra de traiter aisément des questions de convergence des méthodes d'éléments finis. Une métrique permet de mesurer la « distance » entre 2 fonctions. Auparavant, nous définissons le produit scalaire dans un espace fonctionnel qui nous mènera à la notion de métrique et donc de norme.

**Définition 2.12**

Un *produit scalaire* sur un espace fonctionnel linéaire  $S$  est une application définie sur le produit  $S \times S$  qui associe à un couple  $(u, w)$  de  $S \times S$  un scalaire noté  $((u, w))_S$  et vérifiant les propriétés suivantes :

1.  $((u, w))_S = ((w, u))_S \quad \forall u, w \in S$ ;
2.  $((\beta u, w))_S = ((u, \beta w))_S = \beta((u, w))_S \quad \forall \beta \in R \text{ et } \forall u, w \in S$ ;
3.  $((u_1 + u_2, w))_S = ((u_1, w))_S + ((u_2, w))_S \quad \forall u_1, u_2, w \in S$ ;
4.  $((u, u))_S \geq 0$ , le produit scalaire n'étant égal à 0 que si et seulement si  $u \stackrel{\text{p.p.}}{=} 0$ ;

L'espace des fonctions de carré sommable  $L^2(\Omega)$  fournit un premier exemple de produit scalaire. On définit en effet le produit scalaire dans cet espace fonctionnel par :

$$((u, w))_{0, \Omega} = \int_{\Omega} uw \, dv \quad (2.21)$$

et on vérifie relativement facilement les 4 propriétés du produit scalaire. En fait, seule la dernière pose des difficultés puisque  $((u, u))_{0, \Omega}$  peut être nul et ce, même si  $u$  ne s'annule pas en quelques points (plus précisément sur un ensemble de mesure nulle). En pratique, comme nous l'avons déjà mentionné, on ne fait pas de distinction entre 2 fonctions qui ne diffèrent que sur un ensemble de mesure nulle.

**Définition 2.13**

Une *norme* sur un espace fonctionnel linéaire  $S$  est une application de  $S$  dans  $R$  qui associe à une fonction  $u$  de  $S$  un scalaire noté  $\|u\|_S$  et qui vérifie les trois propriétés suivantes :

1. La norme d'une fonction est toujours positive :

$$\|u\|_S > 0, \quad \forall u \in S; \quad \|u\|_S = 0 \Leftrightarrow u \stackrel{\text{p.p.}}{=} 0 \quad (2.22)$$

2. Si  $\beta$  est un scalaire, alors :

$$\|\beta u\|_S = |\beta| \|u\|_S \quad \forall \beta \in R, \quad \forall u \in S \quad (2.23)$$

où  $|\beta|$  est la valeur absolue de  $\beta$ .

3. L'*inégalité triangulaire* :

$$\|u + w\|_S \leq \|u\|_S + \|w\|_S \quad \forall u, w \in S \quad (2.24)$$

Toute application vérifiant ces trois propriétés est une norme. Un espace linéaire peut posséder plusieurs normes. En particulier, si un espace linéaire possède un produit scalaire, il possède automatiquement une norme dite *induite*.

**Définition 2.14**

La *norme induite* sur l'espace  $S$  par le produit scalaire  $((\cdot, \cdot))_S$  est définie par :

$$\|u\|_S = ((u, u))_S^{1/2}$$

Il est facile de vérifier que la norme induite est bel et bien une norme. Nous avons déjà introduit un produit scalaire sur  $L^2(\Omega)$  par la relation 2.21. La norme induite par ce produit scalaire est donc :

$$\|u\|_{0,\Omega} = ((u, u))_{0,\Omega}^{1/2} = \left( \int_{\Omega} u^2 dv \right)^{1/2} \quad (2.25)$$

ce qui semble raisonnable, vu la définition de cet espace. Suivant cette notation, l'inégalité de Cauchy-Schwartz s'écrit plus succinctement sous la forme :

$$|((u, w))_{0,\Omega}| \leq \|u\|_{0,\Omega} \|w\|_{0,\Omega} \quad (2.26)$$

◆ **Remarque 2.8**

Pour en finir avec la notion d'espace fonctionnel linéaire, il nous faudrait parler de la *complétude* d'un espace. Pour ce faire, il faut définir la notion de *suite de Cauchy* et s'assurer que toute suite de Cauchy est convergente dans  $S$ . On dit alors que l'espace  $S$  est *complet*. Encore ici, cela nous entraînerait dans des considérations non essentielles à une bonne compréhension de la suite. ◆

**Définition 2.15**

Un espace fonctionnel linéaire muni d'un produit scalaire (et donc d'une norme induite) et qui est complet est un *espace de Hilbert*.

★ **Théorème 2.3**

L'espace linéaire  $L^2(\Omega)$  muni du produit scalaire 2.21 est un espace de Hilbert. ★

Les espaces de Hilbert jouent un rôle fondamental en éléments finis ainsi que certains espaces de Sobolev que nous introduisons dans la section suivante.

## 2.3 Quelques espaces de Sobolev

Les espaces de Sobolev ont été introduits au début du siècle et ont permis de résoudre bon nombre de problèmes concernant les équations aux dérivées partielles restés sans réponse jusque là. Nous nous limiterons aux espaces les plus utiles en gardant bien à l'esprit que la théorie sous-jacente est beaucoup plus vaste.

Dans ce qui suit, sauf mention explicite du contraire, on suppose l'ouvert  $\Omega$  borné.

### 2.3.1 L'espace $H^1(\Omega)$

#### Définition 2.16

On note  $H^1(\Omega)$  l'espace fonctionnel linéaire défini par :

$$H^1(\Omega) = \left\{ u \in L^2(\Omega) \mid \frac{\partial u}{\partial x_i} \in L^2(\Omega), \quad 1 \leq i \leq 3 \right\}$$

que l'on munit du produit scalaire noté  $((u, w))_{1,\Omega}$  :

$$((u, w))_{1,\Omega} = \int_{\Omega} \left( uw + \sum_{i=1}^3 \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_i} \right) dv$$

et par le fait même d'une norme induite :

$$\|u\|_{1,\Omega} = ((u, u))_{1,\Omega}^{1/2} = \left( \int_{\Omega} \left( u^2 + \sum_{i=1}^3 \left( \frac{\partial u}{\partial x_i} \right)^2 \right) dv \right)^{1/2}$$

#### ◆ Remarque 2.9

Dans la définition de l'espace  $H^1(\Omega)$ , il est important de noter que l'on considère la fonction  $u$  et ses dérivées partielles comme des distributions régulières associées à des fonctions qui sont non seulement dans  $L^1_{loc}(\Omega)$  mais dans  $L^2(\Omega)$  (voir la figure 2.7). Si on dénote  $T_u$  la distribution associée à  $u$ , alors on a :

$$\langle T_u, \phi \rangle = \int_{\Omega} u \phi \, dv \quad \forall \phi \in \mathcal{D}(\Omega)$$

et le terme de droite de cette expression est parfaitement défini puisque  $u \in L^2(\Omega)$ . De plus, si  $u \in H^1(\Omega)$ , il existe des fonctions  $f_i \in L^2(\Omega)$  telles que :

$$\left\langle \frac{\partial T_u}{\partial x_i}, \phi \right\rangle = - \langle T_u, \frac{\partial \phi}{\partial x_i} \rangle = - \int_{\Omega} u \frac{\partial \phi}{\partial x_i} \, dv = \int_{\Omega} f_i \phi \, dv \quad \forall \phi \in \mathcal{D}(\Omega)$$

En effet, si de telles fonctions  $f_i$  existent, alors de la relation 2.16, on a  $\frac{\partial T_u}{\partial x_i} = f_i$  au sens des distributions, ce que l'on note abusivement  $\frac{\partial u}{\partial x_i} = f_i$ . ♦

♦ **Remarque 2.10**

Si une fonction  $u$  est dans  $H^1(\Omega)$ , alors  $u$  et  $\frac{\partial u}{\partial x_i}$  sont automatiquement dans  $L^2(\Omega)$  de sorte que  $H^1(\Omega) \subset L^2(\Omega)$  et que :

$$\|u\|_{1,\Omega}^2 = \|u\|_{0,\Omega}^2 + \sum_{i=1}^3 \left\| \frac{\partial u}{\partial x_i} \right\|_{0,\Omega}^2$$

On a ainsi les inégalités :

$$\|u\|_{0,\Omega} \leq \|u\|_{1,\Omega} \text{ et } \left\| \frac{\partial u}{\partial x_i} \right\|_{0,\Omega} \leq \|u\|_{1,\Omega} \quad (2.27)$$

♦

Le produit scalaire  $((u, w))_{1,\Omega}$  semble naturel car en vertu de l'inégalité de Cauchy-Schwartz 2.26, toutes les intégrales impliquées sont finies et la linéarité de l'intégration permet de vérifier aisément les propriétés d'un produit scalaire.

♦ **Remarque 2.11**

Le produit scalaire et la norme de  $L^2(\Omega)$  ont été notés  $((u, w))_{0,\Omega}$  et  $\|u\|_{0,\Omega}$  par analogie avec l'espace  $H^1(\Omega)$ , puisque l'on peut dire dans ce cas que la fonction et ses dérivées « d'ordre 0 » sont dans  $L^2(\Omega)$ . ♦

Lors de la résolution des équations aux dérivées partielles, il nous faudra introduire des conditions aux limites, ce qui nous amène à parler de la restriction à la frontière  $\Gamma$  d'une fonction de  $H^1(\Omega)$ . Bien que cela semble tout-à-fait naturel de le faire, il faut s'assurer que cette restriction au bord ait un sens du point de vue mathématique. Heureusement, bien que nous ne le justifierons pas de manière tout-à-fait rigoureuse, ces valeurs au bord sont bien définies. Pour s'en convaincre en dimension 1, on peut faire le raisonnement suivant. On suppose que  $\Omega = ]a, b[$  et que  $w \in H^1(]a, b[)$ . On peut donner un sens à la valeur à la frontière  $w(b)$  (et par un argument similaire à  $w(a)$ ) en considérant le fait que l'on a :

$$\left( w(x) \frac{(x-a)}{(b-a)} \right)' = w'(x) \frac{(x-a)}{(b-a)} + w(x) \frac{1}{(b-a)}$$

dont on déduit, en intégrant sur  $]a, b[$  que :

$$w(b) = \int_a^b w'(x) \frac{(x-a)}{(b-a)} dx + \int_a^b w(x) \frac{1}{(b-a)} dx$$

Chacun des termes de droite de cette équation est bien défini en vertu de l'inégalité de Cauchy-Schwartz car  $w(x)$  et  $w'(x)$  sont dans  $L^2(]a, b[)$ . Il en ressort de plus, toujours en utilisant l'inégalité de Cauchy-Schwartz que :

$$|w(b)| \leq \|w'(x)\|_{0,\Omega} \left\| \frac{(x-a)}{(b-a)} \right\|_{0,\Omega} + \|w(x)\|_{0,\Omega} \left\| \frac{1}{(b-a)} \right\|_{0,\Omega} \leq C \|w\|_{1,\Omega} \quad (2.28)$$

Dans cette inégalité, la constante  $C$  ne dépend que de la géométrie c.-à-d. de  $a$  et  $b$ .

◆ **Remarque 2.12**

Le raisonnement précédent n'est plus valide si la fonction  $w$  n'est que dans  $L^2(\Omega)$ . On ne peut donc pas parler de la valeur au bord (sur la frontière de  $\Omega$ ) d'une fonction de  $L^2(\Omega)$  autrement que dans un sens très faible. ◆

**Définition 2.17**

La restriction au bord  $\Gamma$  d'une fonction de  $w \in H^1(\Omega)$  est appelée *trace au bord* de  $w$  et est notée  $w|_{\Gamma}$  ou encore  $\gamma_0(w)$ .

★ **Théorème 2.4**

L'ensemble des traces au bord des fonctions de  $H^1(\Omega)$  forme un sous-espace de  $L^2(\Gamma)$  noté  $H^{1/2}(\Gamma)$ . Plus succinctement, on a :

$$\gamma_0(H^1(\Omega)) = H^{1/2}(\Gamma) \subsetneq L^2(\Gamma) \quad (2.29)$$

**Démonstration** : voir Reddy [30] ou Raviart-Thomas [29]. ★

◆ **Remarque 2.13**

De cette définition, on conclut immédiatement que si  $g \in H^{1/2}(\Gamma)$ , alors il existe forcément au moins une fonction  $w_g \in H^1(\Omega)$  dont  $g$  est la trace au bord c.-à-d. :

$$\gamma_0(w_g) = w_g|_{\Gamma} = g \quad (2.30)$$

Cette remarque sera essentielle lors de l'imposition des conditions aux limites dans les équations aux dérivées partielles d'ordre 2. Dans le cadre d'une méthode d'éléments finis, pour une condition aux limites  $g$  donnée, nous construirons explicitement la fonction  $w_g$  de  $H^1(\Omega)$ . La fonction  $w_g$  sera appelée le relèvement de la condition aux limites  $g$ . ◆

Le raisonnement qui nous a mené à l'inégalité 2.28 peut être généralisé à la frontière d'un domaine quelconque en plusieurs dimensions.

★ **Théorème 2.5 (Continuité de la trace au bord)**

Si  $w \in H^1(\Omega)$ , il existe une constante  $C$  telle que :

$$\|\gamma_0(w)\|_{0,\Gamma} \leq C \|w\|_{1,\Omega} \quad (2.31)$$

où :

$$\|\gamma_0(w)\|_{0,\Gamma} = \left( \int_{\Gamma} (\gamma_0(w))^2 ds \right)^{1/2}$$

Le résultat est également vrai si on remplace  $\Gamma$  par une partie non nulle  $\Gamma_0$  de  $\Gamma$ . En dimension 1, on remplace tout simplement la norme  $\|\cdot\|_{0,\Gamma}$  par la valeur absolue comme dans l'équation 2.28. ★

Ce dernier résultat est connu sous le nom de *continuité de la trace au bord* et généralise l'équation 2.28. Il signifie simplement que la trace au bord d'une fonction  $w$  de  $H^1(\Omega)$  dépend continûment de la fonction  $w$ . Une autre manière de voir les choses consiste à dire que si  $w$  tend vers 0, alors sa trace au bord tend également vers 0.

**Définition 2.18**

On définit l'espace  $H_0^1(\Omega)$  comme la fermeture de  $\mathcal{D}(\Omega)$  pour la norme  $\|\cdot\|_{1,\Omega}$ . Ainsi, pour chaque  $v \in H_0^1(\Omega)$ , il existe une suite  $v_i$  de fonctions de  $\mathcal{D}(\Omega)$  et telle que :

$$\lim_{i \rightarrow \infty} \|v - v_i\|_{1,\Omega} = 0$$

Les fonctions de  $H_0^1(\Omega)$  s'annulent donc au bord et on peut écrire :

$$H_0^1(\Omega) = \{w \in H^1(\Omega) \mid w = 0 \text{ sur } \Gamma\}$$

On peut aussi définir :

$$H_{\Gamma_0}^1(\Omega) = \{w \in H^1(\Omega) \mid w = 0 \text{ sur } \Gamma_0\}$$

où  $\Gamma_0$  est une partie de la frontière  $\Gamma$  du domaine  $\Omega$ . En dimension 1, on a par exemple :

$$H^1(]a, b]) = \{w(x) \in L^2(]a, b]) \mid w'(x) \in L^2(]a, b])\}$$

et

$$H_0^1(]a, b]) = \{w \in H^1(]a, b]) \mid w(a) = w(b) = 0\}$$

tandis que l'on pourrait poser :

$$H_{\Gamma_0}^1(]a, b]) = \{w \in H^1(]a, b]) \mid w(a) = 0\}$$

ou encore :

$$H_{\Gamma_0}^1(]a, b[) = \{w \in H^1(]a, b[) \mid w(b) = 0\}$$

► **Exemple 2.15**

Choisissons encore  $\Omega = ]0, 1[$ . La fonction  $f(x) = x^{3/4}$  est dans  $H^1(]0, 1[)$ . En effet :

$$\int_0^1 (x^{3/4})^2 dx = \int_0^1 (x^{3/2}) dx = \frac{2}{5} < \infty$$

De plus,  $f'(x) = (3/4)x^{-1/4}$  et on a :

$$\int_0^1 (f'(x))^2 dx = \frac{9}{16} \int_0^1 x^{-1/2} dx = \frac{9}{8} < \infty$$

Il convient toutefois de remarquer que la fonction  $f'(x)$  n'est pas définie en  $x = 0$ . ◀

► **Exemple 2.16**

Choisissons cette fois  $\Omega = ]-1, 1[$  et la fonction de Heaviside (voir la figure 2.4) restreinte à cet intervalle. Il est facile de montrer que  $H(x) \in L^2(]-1, 1[)$ . Par contre, nous avons vu qu'au sens des distributions,  $T'_H = \delta_0$  et que la distribution de Dirac n'est pas régulière. On ne peut donc pas associer à cette distribution une fonction de  $L^1_{loc}(]-1, 1[)$  et encore moins  $L^2(]-1, 1[)$ . La fonction de Heaviside n'est donc pas dans  $H^1(]-1, 1[)$ . Il en est de même, en une dimension, de toutes les fonctions qui possèdent une discontinuité de première espèce dans  $\Omega$ . ◀

★ **Théorème 2.6**

L'expression :

$$|u|_{1,\Omega} = \left( \sum_{i=1}^n \int_{\Omega} \left( \frac{\partial u}{\partial x_i} \right)^2 dv \right)^{1/2} \quad (2.32)$$

est une norme sur les espaces  $H_0^1(\Omega)$  et  $H_{\Gamma_0}^1(\Omega)$ .

**Esquisse de la démonstration :**

Il faut tout d'abord remarquer la différence entre cette norme (notation avec une seule barre) et la norme habituelle sur  $H^1(\Omega)$  (notation avec 2 barres). C'est l'absence du terme en  $u^2$ . Il nous faut donc démontrer les 3 propriétés d'une norme. La seule propriété qui pose des difficultés est la première qui requiert qu'une norme ne s'annule que pour la fonction  $u \stackrel{\text{p.p.}}{=} 0$ . Il est en effet évident que la quantité 2.32 est toujours positive ou nulle. De plus, si  $|u|_{1,\Omega} = 0$ , on peut immédiatement conclure que toutes les dérivées partielles de  $u$  sont nulles (p.p.) et donc que  $u$  est une constante. La conclusion vient du fait que puisque  $u$  est nulle au bord (ou sur une partie  $\Gamma_0$  du bord), cette constante est forcément 0. ★

◆ **Remarque 2.14**

Il est extrêmement important de noter que la quantité 2.32 n'est pas une norme sur  $H^1(\Omega)$ . Le raisonnement précédent ne peut en effet s'appliquer puisque les fonctions de  $H^1(\Omega)$  ne s'annulent pas toutes au bord. ◆

**Définition 2.19**

Deux normes  $\|\cdot\|_1$  et  $\|\cdot\|_2$  sur un espace de Hilbert  $V$  sont dites *équivalentes* s'il existe des constantes  $C_1$  et  $C_2$  telles que :

$$C_1 \|w\|_2 \leq \|w\|_1 \leq C_2 \|w\|_2 \quad \forall w \in V \quad (2.33)$$

★ **Théorème 2.7 (Inégalité de Poincaré)**

Les normes  $\|\cdot\|_{1,\Omega}$  et  $|\cdot|_{1,\Omega}$  sont des normes équivalentes sur les espaces  $H_0^1(\Omega)$  et  $H_{\Gamma_0}^1(\Omega)$  (mais pas sur  $H^1(\Omega)$ ).

**Démonstration** (voir Raviart-Thomas [29]). ★

★ **Théorème 2.8 (Inégalité de Poincaré)**

Soit  $\Omega$  un ouvert de  $\mathcal{R}^n$  borné dans au moins une direction. Il existe alors une constante  $C$  qui ne dépend que de  $\Omega$  et telle que :

$$\|v\|_{0,\Omega} \leq C(\Omega) |v|_{1,\Omega} \quad \forall v \in H_0^1(\Omega) \quad (2.34)$$

**Démonstration :**

Nous reprenons ici la démonstration de Lucquin [24]. Par définition de  $H_0^1(\Omega)$ , il suffit de démontrer le résultat pour  $v \in \mathcal{D}(\Omega)$  et comme ces dernières fonctions s'annulent au voisinage de la frontière de  $\Omega$ , on peut les prolonger par 0 dans tout  $\mathcal{R}^n$ . On supposera pour fixer les idées que  $\Omega$  est borné dans la direction  $x_1$  et donc que la composante  $x_1$  du support des fonctions  $v$  est inclus dans un intervalle  $]a, b[$ . On a alors :

$$v(x_1, \mathbf{x}') = \int_a^{x_1} \frac{\partial v}{\partial x_1}(u, \mathbf{x}') du$$

car  $v(a, \mathbf{x}') = 0$ . L'inégalité de Cauchy-Schwarz nous donne alors :

$$|v(x_1, \mathbf{x}')|^2 \leq (x_1 - a) \int_a^{x_1} \left| \frac{\partial v}{\partial x_1}(u, \mathbf{x}') \right|^2 du \leq (b - a) \int_a^b \left| \frac{\partial v}{\partial x_1}(u, \mathbf{x}') \right|^2 du$$

Il reste à intégrer sur les autres variables d'espace, soit sur  $\mathbf{x}'$ . On a ainsi, en vertu de la dernière inégalité :

$$\int_{\mathcal{R}^{n-1}} \left| \frac{\partial v}{\partial x_1}(u, \mathbf{x}') \right|^2 d\mathbf{x}' \leq (b-a) \left\| \frac{\partial v}{\partial x_1} \right\|_{0, \mathcal{R}^n}^2$$

Une dernière intégration par rapport à  $x_1$  donne :

$$\|v\|_{0, \mathcal{R}^n} \leq (b-a)^2 \left\| \frac{\partial v}{\partial x_1} \right\|_{0, \mathcal{R}^n} \leq (b-a)^2 |v|_{1, \mathcal{R}^n}$$

et le résultat suit car  $\|v\|_{0, \mathcal{R}^n} = \|v\|_{0, \Omega}$  et  $\left\| \frac{\partial v}{\partial x_1} \right\|_{0, \mathcal{R}^n} = \left\| \frac{\partial v}{\partial x_1} \right\|_{0, \Omega}$  par prolongement par 0.

Les normes  $\|\cdot\|_{1, \Omega}$  et  $|\cdot|_{1, \Omega}$  sont donc des normes équivalentes sur les espaces  $H_0^1(\Omega)$  et un résultat similaire existe pour  $H_{\Gamma_0}^1(\Omega)$  (voir [24]). Le résultat est en général faux pour  $H^1(\Omega)$ . ★

### 2.3.2 L'espace $H^2(\Omega)$

#### Définition 2.20

On note  $H^2(\Omega)$  l'espace défini par :

$$H^2(\Omega) = \left\{ u \in L^2(\Omega) \left| \frac{\partial u}{\partial x_i} \in L^2(\Omega), \frac{\partial^2 u}{\partial x_i \partial x_j} \in L^2(\Omega) \quad 1 \leq i, j \leq 3 \right. \right\}$$

que l'on munit du produit scalaire noté  $((u, w))_{2, \Omega}$  :

$$((u, w))_{2, \Omega} = \int_{\Omega} \left( uw + \sum_{i=1}^3 \left( \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_i} \right) + \sum_{i,j=1}^3 \left( \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{\partial^2 w}{\partial x_i \partial x_j} \right) \right) dv$$

et de la norme induite :

$$\|u\|_{2, \Omega} = \left( \int_{\Omega} \left( u^2 + \sum_{i=1}^3 \left( \frac{\partial u}{\partial x_i} \right)^2 + \sum_{i,j=1}^3 \left( \frac{\partial^2 u}{\partial x_i \partial x_j} \right)^2 \right) dv \right)^{1/2}$$

Les dérivées sont ici encore prises au sens des distributions. On a aussi :

$$\|u\|_{2, \Omega}^2 = \|u\|_{0, \Omega}^2 + \sum_{i=1}^3 \left\| \frac{\partial u}{\partial x_i} \right\|_{0, \Omega}^2 + \sum_{i,j=1}^3 \left\| \frac{\partial^2 u}{\partial x_i \partial x_j} \right\|_{0, \Omega}^2 = \|u\|_{1, \Omega}^2 + \sum_{i,j=1}^3 \left\| \frac{\partial^2 u}{\partial x_i \partial x_j} \right\|_{0, \Omega}^2$$

Puisque l'on requiert que les dérivées d'ordre 2 soit dans  $L^2(\Omega)$ , les fonctions de  $H^2(\Omega)$  sont plus régulières que celles de  $H^1(\Omega)$ . Il en est de même pour la trace au bord.

**Définition 2.21**

Si  $w \in H^2(\Omega)$ , alors la fonction  $\partial w / \partial \mathbf{n}$  définie sur la frontière  $\Gamma$  est appelée la *trace normale* de  $w$  au bord que l'on note  $\gamma_1(w)$ . Rappelons que :

$$\gamma_1(w) = \frac{\partial w}{\partial \mathbf{n}} = \nabla w \cdot \mathbf{n}$$

où  $\mathbf{n}$  est le vecteur normal unitaire à  $\Gamma$ .

★ **Théorème 2.9**

L'ensemble des traces au bord des fonctions de  $H^2(\Omega)$  forme un sous-espace de  $L^2(\Gamma)$  noté  $H^{3/2}(\Gamma)$ . Plus succinctement, on a :

$$\gamma_0(H^2(\Omega)) = H^{3/2}(\Gamma) \subsetneq H^{1/2}(\Gamma) \subsetneq L^2(\Gamma) \quad (2.35)$$

De plus, l'ensemble des traces normales au bord des fonctions de  $H^2(\Omega)$  forme un sous-espace de  $L^2(\Gamma)$  qui n'est autre que  $H^{1/2}(\Gamma)$ . On a :

$$\gamma_1(H^2(\Omega)) = H^{1/2}(\Gamma) \subsetneq L^2(\Gamma)$$

**Démonstration** (voir Raviart-Thomas [29]). ★

★ **Théorème 2.10**

Si  $g \in H^{3/2}(\Gamma)$  et  $h \in H^{1/2}(\Gamma)$ , alors il existe au moins une fonction  $w_{gh} \in H^2(\Omega)$  (dite de relèvement des conditions aux limites) dont  $g$  est la trace au bord c.-à-d. :

$$\gamma_0(w_{gh}) = w_{gh}|_{\Gamma} = g$$

et dont  $h$  est la trace normale au bord c.-à-d.

$$\gamma_1(w_{gh}) = \frac{\partial w_{gh}}{\partial \mathbf{n}} \Big|_{\Gamma} = h$$

De plus, on a la continuité de la trace au bord c.-à-d. :

$$\|\gamma_0(w)\|_{0,\Gamma} + \|\gamma_1(w)\|_{0,\Gamma} = \|g\|_{0,\Gamma} + \|h\|_{0,\Gamma} \leq C \|w\|_{2,\Omega} \quad (2.36)$$

En dimension 1, on remplace tout simplement la norme  $\|\cdot\|_{0,\Gamma}$  par la valeur absolue.

**Démonstration** (voir Raviart-Thomas [29]). ★

◆ **Remarque 2.15**

Ce résultat sera essentiel lors de l'imposition des conditions aux limites dans les équations aux dérivées partielles d'ordre 4. Nous supposons de plus que le résultat est vrai lorsque les conditions aux limites sur  $w$  et  $\partial w/\partial \mathbf{n}$  ne sont données que sur des parties de la frontière  $\Gamma$ . ◆

**Définition 2.22**

L'espace  $H_0^2(\Omega)$  est défini par :

$$H_0^2(\Omega) = \left\{ w \in H^2(\Omega) \mid w = \frac{\partial w}{\partial \mathbf{n}} = 0 \text{ sur } \Gamma \right\}$$

Il est de toute évidence possible de construire des variantes des espaces précédents. Il s'agit de sous-espaces de  $H^2(\Omega)$  pour lesquels la fonction  $w$  et sa trace normale s'annulent seulement sur certaines parties de la frontière. Par exemple, on pourrait prendre :

$$\left\{ w \in H^2(\Omega) \mid w = 0 \text{ sur } \Gamma_0, \frac{\partial w}{\partial \mathbf{n}} = 0 \text{ sur } \Gamma_1 \right\}$$

où  $\Gamma_0$  et  $\Gamma_1$  sont des parties de la frontière  $\Gamma$  disjointes ou non et qui ne recouvrent pas forcément toute la frontière  $\Gamma$ .

★ **Théorème 2.11**

L'expression :

$$|u|_{2,\Omega} = \left( \int_{\Omega} (\nabla^2 u)^2 dv \right)^{1/2} \quad (2.37)$$

est une norme sur l'espace  $H_0^2(\Omega)$ .

**Démonstration** (voir Ciarlet [9]). ★

◆ **Remarque 2.16**

Il est important de remarquer que tout comme la norme 2.32 n'est pas une norme sur  $H^1(\Omega)$ , la quantité 2.37 n'est pas une norme sur  $H^2(\Omega)$ . Le résultat précédent peut également s'étendre au cas où la fonction et sa trace normale ne s'annulent que sur des parties  $\Gamma_0$  et  $\Gamma_1$  respectivement. Cette généralisation est cependant assez délicate. ◆

La situation est plus simple en dimension 1 et on a par exemple :

$$H^2(]a, b[) = \{w(x) \in L^2(]a, b[) \mid w'(x) \in L^2(]a, b[), w''(x) \in L^2(]a, b[)\}$$

ainsi que :

$$H_0^2(]a, b[) = \{w(x) \in H^2(]a, b[) \mid w(a) = w(b) = w'(a) = w'(b) = 0\}$$

La norme 2.37 se réduit dans ce cas à :

$$\|u\|_{2,]a,b[} = \left( \int_a^b (u''(x))^2 dx \right)^{1/2} \quad (2.38)$$

### ★ Théorème 2.12

La quantité 2.38 est une norme équivalente à la norme  $\|\cdot\|_{2,\Omega}$  sur l'espace  $H_0^2(]a, b[)$  ainsi que sur les espaces suivants :

- $V_1 = \{w \in H^2(]a, b[) \mid w(a) = w(b) = 0\}$
- $V_2 = \{w \in H^2(]a, b[) \mid w(a) = w'(b) = 0\}$
- $V_3 = \{w \in H^2(]a, b[) \mid w'(a) = w(b) = 0\}$
- $V_4 = \{w \in H^2(]a, b[) \mid w(a) = w'(a) = 0\}$
- $V_5 = \{w \in H^2(]a, b[) \mid w(b) = w'(b) = 0\}$

mais pas sur l'espace :

$$X = \{w(x) \in H^2(]a, b[) \mid w'(a) = w'(b) = 0\}$$

### Esquisse de la démonstration :

Les deux premières propriétés d'une norme ne posent aucune difficulté particulière. Par contre, si  $\|u\|_{2,]a,b[} = 0$ , on en déduit que  $u''(x) = 0$  (presque partout!) et donc que  $u(x) = cx + d$ . Les conditions aux limites permettent alors de montrer que  $u(x) = 0$  lorsque  $u(x) \in V_1, V_2, V_3, V_4$  ou  $V_5$ . Par contre, si  $u(x) \in X$ , on ne peut que conclure que  $u(x) = d$ . ★

### ◆ Remarque 2.17

On peut de manière générale définir les espaces  $H^m(\Omega)$  en ajoutant les dérivées partielles jusqu'à l'ordre  $m$ . ◆

## 2.4 Un résultat d'approximation

La méthode des éléments finis utilise des sous-domaines (les éléments) de forme géométrique simple (des intervalles en dimension 1, des triangles ou des quadrilatères en dimension 2, des tétraèdres en dimension 3) sur lesquels on construit des approximations polynômiales. En d'autres

termes, la solution approximative de l'équation différentielle de départ est constituée de polynômes différents sur chaque sous-domaine. Il est tout de même important de s'assurer que l'approximation ainsi obtenue soit dans un espace de Sobolev approprié. C'est l'objet du résultat suivant.

★ **Théorème 2.13**

Soit un domaine  $\Omega$  constitué de sous-domaines  $\Omega_i$  et une fonction  $u(x)$  telle que sa restriction à chaque sous-domaine  $\Omega_i$  est un polynôme de degré  $n$ . On suppose de plus que les polynômes de degré  $n$  sont dans  $H^1(\Omega_i)$  quel que soit  $i$ . Alors si la fonction  $u(x)$  est continue à la frontière entre les sous-domaines  $\Omega_i$ , la fonction  $u(x)$  appartient à l'espace  $H^1(\Omega)$ . Si de plus, la fonction  $u(x)$  est différentiable au sens classique c.-à-d. ses dérivées partielles d'ordre 1 sont continues à la frontière des sous-domaines, alors  $u(x) \in H^2(\Omega)$ .

**Démonstration** (voir Ciarlet [9]) :

Ce résultat est évident en dimension 1. En effet, si la fonction  $u(x)$  possède une discontinuité de première espèce à la frontière de 2 sous-intervalles, alors sa dérivée fait intervenir une distribution de Dirac qui n'est pas dans  $L^2(\Omega)$ .

Dans le cas général, il est clair que la fonction  $u$  est dans  $L^2(\Omega)$  puisqu'il suffit de sommer sur les sous-domaines. Ce qui est moins évident c'est de vérifier que les dérivées partielles de  $u$  sont dans  $L^2(\Omega)$  au sens des distributions. Pour le montrer, il nous faut trouver des fonctions  $f_j$  de  $L^2(\Omega)$  telles que :

$$\int_{\Omega} f_j \phi \, dv = - \int_{\Omega} u \frac{\partial \phi}{\partial x_j} \, dv \quad \forall \phi \in \mathcal{D}(\Omega)$$

Si de telles fonctions existent, en vertu de la définition 2.16, on a  $\frac{\partial u}{\partial x_j} = f_j$  au sens des distributions et les dérivées partielles sont donc dans  $L^2(\Omega)$ . Il semble naturel de considérer comme fonctions  $f_j$  celles constituées des dérivées partielles de la fonction  $u$  restreintes à chaque sous-domaine  $\Omega_i$  c.-à-d.

$$f_j|_{\Omega_i} = \left. \frac{\partial u}{\partial x_j} \right|_{\Omega_i}$$

qui sont bien définies sur chaque sous-domaine puisque les polynômes de degré  $n$  sont dans  $H^1(\Omega_i)$ . On a alors en intégrant par parties (voir la relation A.4) :

$$\int_{\Omega_i} f_j|_{\Omega_i} \phi \, dv = \int_{\Omega_i} \left. \frac{\partial u}{\partial x_j} \right|_{\Omega_i} \phi \, dv = - \int_{\Omega_i} u|_{\Omega_i} \frac{\partial \phi}{\partial x_j} \, dv + \int_{\partial \Omega_i} u|_{\Omega_i} \phi \, n_{ij} \, ds$$

où  $n_{ij}$  est la  $j^e$  composante du vecteur normal  $\mathbf{n}_i$  à la frontière de  $\Omega_i$  (voir la relation A.4). En additionnant les contributions de chaque sous-domaine, on trouve :

$$\int_{\Omega} f_j \phi \, dv = - \int_{\Omega} u \frac{\partial \phi}{\partial x_j} \, dv + \sum_i \int_{\partial \Omega_i} u|_{\Omega_i} \phi \, n_{ij} \, ds$$

Le dernier terme de droite est une somme sur les frontières des sous-domaines. D'une part, si  $\partial \Omega_i \subset \partial \Omega$ , ce terme est nul puisque les fonctions  $\phi$  y sont nulles. À la frontière commune entre 2

sous-domaines,  $\partial\Omega_i$  et  $\partial\Omega_k$ , on a :

$$\int_{\partial\Omega_i} u|_{\Omega_i} \phi n_{ij} ds + \int_{\partial\Omega_k} u|_{\Omega_k} \phi n_{kj} ds$$

Ces termes s'annulent deux à deux en vertu de la continuité de la fonction  $u$  et du changement de signe du vecteur normal ( $n_{ij} = -n_{kj}$ ). Les dérivées partielles sont donc dans  $L^2(\Omega)$  et le théorème est démontré. On peut effectuer un raisonnement similaire pour démontrer que la fonction  $u$  est dans  $H^2(\Omega)$  si  $u$  est non seulement continue mais différentiable à la frontière des sous-domaines. ★

◆ **Remarque 2.18**

Ce résultat est fondamental car il sera à la base des techniques d'approximation par éléments finis où les sous-domaines  $\Omega_i$  seront tout simplement les éléments. ◆

## 2.5 Exercices

1. Calculer les dérivées premières et secondes au sens des distributions des fonctions suivantes :

a)

$$f(x) = \begin{cases} 1 & \text{si } 0 < x < 1 \\ 0 & \text{ailleurs} \end{cases}$$

b)

$$f(x) = \begin{cases} -x & \text{si } x < 0 \\ \sin(x) & \text{si } x > 0 \end{cases}$$

c)

$$f(x) = \begin{cases} -x & \text{si } x < 0 \\ \cos(x) & \text{si } x > 0 \end{cases}$$

2. Les fonctions de l'exercice précédent sont-elles dans  $L^2(]-1, 1[)$  ? dans  $H^1(]-1, 1[)$  ? dans  $H^2(]-1, 1[)$  ?
3. Calculer, au sens des distributions,  $\nabla \cdot (k\nabla u)$  dans le carré  $]0, 1[ \times ]0, 1[$  où :

$$u(x, y) = \begin{cases} 2(y-x) & \text{si } x < y \\ x^2 - y^2 & \text{si } x > y \end{cases} \quad \text{et } k = \begin{cases} \sqrt{2} & \text{si } x < y \\ \frac{\sqrt{2}}{2} & \text{si } x > y \end{cases}$$

**Suggestion :** Calculer d'abord  $\nabla u$  et bien regarder ce qui se passe sur l'axe  $y = x$ .

4. Si  $\Omega = ]-\infty, \infty[$ , la fonction  $f(x) = 1$  est-elle dans l'espace des fonctions localement intégrables  $L^1_{loc}(\Omega)$ . Même question pour la fonction  $f(x) = 1/x$  dans  $\Omega = ]0, 1[$ .
5. La fonction  $f(x) = |x|$  est-elle dans  $H^1(]-1, 1[)$  ? dans  $H^2(]-1, 1[)$  ?
6. La fonction  $f(x) = x(1-x)$  est-elle dans  $H^2_0(]0, 1[)$  ? Même question pour la fonction  $f(x) = x^2(1-x)^2$  ?
7. Pour quelles valeurs réelles de  $r$ , la fonction  $x^r$  est-elle dans  $L^2(]0, 1[)$  ? dans  $H^1(]0, 1[)$  ? dans  $H^2(]0, 1[)$  ?

## Chapitre 3

# Théorème de Lax-Milgram

Comme nous l'avons vu au chapitre 1, la résolution d'une équation aux dérivées partielles se ramène à celle d'un problème variationnel. Nous abordons dans ce chapitre les outils de base pour l'étude des formulations variationnelles.

### 3.1 Formes linéaires et bilinéaires

**Définition 3.1**

On appelle *forme linéaire* une fonctionnelle linéaire sur un espace de Hilbert  $V$ . Une forme linéaire  $l$  vérifie donc les propriétés suivantes :

- $l(\beta w) = \beta l(w) \quad \forall w \in V \text{ et } \forall \beta \in \mathbb{R}$
- $l(w_1 + w_2) = l(w_1) + l(w_2) \quad \forall w_1, w_2 \in V$

**► Exemple 3.1**

Si  $V = L^2(\Omega)$  et  $f(\mathbf{x}) \in L^2(\Omega)$  alors :

$$l_f(w) = \int_{\Omega} f(\mathbf{x})w(\mathbf{x}) \, dv$$

est une forme linéaire sur  $V$  comme nous l'avons démontré au chapitre précédent. ◀

**Définition 3.2**

Une forme linéaire  $l$  sur l'espace de Hilbert  $V$  muni de la norme  $\|\cdot\|_V$ , est dite *continue* s'il existe une constante  $C$  telle que :

$$|l(w)| \leq C \|w\|_V \quad \forall w \in V \tag{3.1}$$

► **Exemple 3.2**

La fonctionnelle  $l_f$  de l'exemple précédent est continue. En effet, de l'inégalité de Cauchy, on a :

$$|l_f(w)| = \left| \int_{\Omega} f(\mathbf{x})w(\mathbf{x}) \, dv \right| \leq \|f\|_{0,\Omega} \|w\|_{0,\Omega}$$

et l'inégalité 3.1 suit en posant  $C = \|f\|_{0,\Omega}$ . ◀

**Définition 3.3**

L'ensemble de toutes les formes linéaires continues sur un espace de Hilbert  $V$  est appelé *espace dual* de  $V$  et est noté  $V'$ .

◆ **Remarque 3.1**

Nous avons déjà rencontré la notion de dualité lorsque nous avons défini l'espace des distributions  $\mathcal{D}'(\Omega)$  qui n'est cependant pas un espace de Hilbert. Les duaux des espaces  $H_0^1(\Omega)$  et  $H_0^2(\Omega)$  sont notés respectivement  $H^{-1}(\Omega)$  et  $H^{-2}(\Omega)$ . Notez bien que  $H^{-1}(\Omega)$  est le dual de  $H_0^1(\Omega)$  et non de  $H^1(\Omega)$ . ◆

**Définition 3.4**

Une *forme bilinéaire* sur un espace de Hilbert  $V$  est une application  $a$  qui associe à un couple  $(u, w) \in V \times V$  un scalaire noté  $a(u, w)$  satisfaisant :

$$- a(\beta_1 u_1 + \beta_2 u_2, w) = \beta_1 a(u_1, w) + \beta_2 a(u_2, w) \quad \forall u_1, u_2, w \in V \text{ et } \forall \beta_1, \beta_2 \in \mathbb{R}$$

$$- a(u, \beta_1 w_1 + \beta_2 w_2) = \beta_1 a(u, w_1) + \beta_2 a(u, w_2) \quad \forall u, w_1, w_2 \in V \text{ et } \forall \beta_1, \beta_2 \in \mathbb{R}$$

Une forme bilinéaire est donc linéaire en chacun de ses 2 arguments.

**Définition 3.5**

Une forme bilinéaire  $a$  est dite *continue* sur  $V \times V$  s'il existe une constante  $C$  telle que :

$$|a(u, w)| \leq C \|u\|_V \|w\|_V \quad \forall u, w \in V \quad (3.2)$$

**Définition 3.6**

Une forme bilinéaire  $a$  est dite *symétrique* si :

$$a(u, w) = a(w, u) \quad \forall u, w \in V$$

**► Exemple 3.3**

L'application  $a$  définie par :

$$a(u, w) = \int_{\Omega} uw \, dv$$

est une forme bilinéaire symétrique et continue sur  $L^2(\Omega)$ . La bilinéarité est en effet triviale à démontrer. De plus, la continuité vient encore une fois de l'inégalité de Cauchy :

$$\left| \int_{\Omega} uw \, dv \right| \leq \|u\|_{0,\Omega} \|w\|_{0,\Omega}$$

et il ne reste qu'à choisir la constante  $C$  égale à 1 dans l'inégalité 3.2. ◀

**► Exemple 3.4**

L'application  $a$  définie par :

$$a(u, w) = \int_{\Omega} (uw + \nabla u \cdot \nabla w) \, dv = \int_{\Omega} \left( uw + \sum_{i=1}^3 \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_i} \right) \, dv$$

est une forme bilinéaire continue sur  $H^1(\Omega)$ . La bilinéarité est encore ici facile à vérifier et la continuité découle de l'inégalité de Cauchy :

$$\begin{aligned} \left| \int_{\Omega} uw + \nabla u \cdot \nabla w \, dv \right| &\leq \|u\|_{0,\Omega} \|w\|_{0,\Omega} + \sum_{i=1}^3 \left\| \frac{\partial u}{\partial x_i} \right\|_{0,\Omega} \left\| \frac{\partial w}{\partial x_i} \right\|_{0,\Omega} \\ &\leq 4\|u\|_{1,\Omega} \|w\|_{1,\Omega} \end{aligned}$$

Dans la dernière inégalité, nous avons fait usage des inégalités 2.27. Il ne reste qu'à choisir la constante  $C$  égale à 4 dans l'inégalité 3.2. ◀

**Définition 3.7**

Une forme bilinéaire est dite *coercive* ou *elliptique* s'il existe une constante strictement positive  $\alpha$  telle que :

$$a(w, w) \geq \alpha \|w\|_V^2 \quad \forall w \in V \quad (3.3)$$

Une forme bilinéaire coercive est une généralisation de la notion de *matrice définie positive* que nous reverrons un peu plus loin.

**3.2 Théorème de Lax-Milgram**

Nous en arrivons au résultat le plus fondamental de cette section. Beaucoup de formulations variationnelles entrent dans le cadre de ce théorème.

**★ Théorème 3.1 (de Lax-Milgram)**

Soit  $V$  un espace de Hilbert et soit  $l$  et  $a$  des formes linéaire et bilinéaire continues sur  $V$  et  $V \times V$  respectivement ( $l \in V'$ ). Si de plus  $a$  est coercive, alors il existe une unique solution  $u$  du problème variationnel :

♠ trouver une fonction  $u \in V$  telle que :

$$a(u, w) = l(w) \quad \forall w \in V \quad (3.4)$$

**Démonstration** (facultative) :

Nous ne ferons la démonstration que dans le cas où la forme bilinéaire est symétrique. La forme bilinéaire étant coercive, elle satisfait les propriétés d'un produit scalaire sur  $V$  et la norme induite par ce produit scalaire est équivalente à  $\|\cdot\|_V$ . L'espace  $V$  muni de ce produit scalaire est donc un espace de Hilbert. Du théorème de représentation de Riesz, il existe  $u \in V$  (qui dépend de  $l$ ) tel que :

$$a(u, w) = l(w) \quad \forall w \in V$$

On a donc l'existence. En ce qui concerne l'unicité, s'il existait deux solutions  $u_1$  et  $u_2$ , alors en soustrayant, on aurait :

$$a(u_1 - u_2, w) = l(w) - l(w) = 0 \quad \forall w \in V$$

et en particulier, en prenant  $w = u_1 - u_2$ , la coercivité entraîne que  $u_1 = u_2$ , ce qui complète la démonstration. On trouvera la généralisation de ce résultat dans Ciarlet [9]. ★

**★ Théorème 3.2**

Sous les mêmes hypothèses que celles du théorème de Lax-Milgram et si de plus la forme bilinéaire  $a$  est symétrique, le problème variationnel 3.4 est équivalent au problème de minimisation suivant :

♠ trouver  $u \in V$  telle que :

$$J(u) = \inf_{w \in V} J(w) = \inf_{w \in V} \frac{1}{2}a(w, w) - l(w) \quad (3.5)$$

La fonctionnelle  $J$  est souvent appelée *fonctionnelle d'énergie* dans les applications.

**Démonstration :**

En premier lieu, on démontre que si  $u$  est l'unique solution du problème variationnel 3.4 alors  $u$  minimise forcément la fonctionnelle  $J$  sur tout l'espace  $V$ . Soit donc  $w \in V$  quelconque. On peut toujours écrire que :

$$w = u + (w - u) = u + w_1 \quad (w_1 = w - u)$$

On a alors :

$$J(w) = J(u + w_1) = \frac{1}{2}a(u + w_1, u + w_1) - l(u + w_1)$$

Puisque  $a$  est bilinéaire et  $l$  linéaire, on a :

$$J(w) = \frac{1}{2} [a(u, u) + a(u, w_1) + a(w_1, u) + a(w_1, w_1)] - l(u) - l(w_1)$$

et la symétrie de  $a$  nous donne :

$$J(w) = \left( \frac{1}{2}a(u, u) - l(u) \right) + (a(u, w_1) - l(w_1)) + \frac{1}{2}a(w_1, w_1)$$

Dans le terme de droite, on reconnaît dans la première parenthèse  $J(u)$  tandis que l'expression à l'intérieur de la deuxième parenthèse est nulle puisque  $u$  est la solution du problème 3.4. Enfin, la coercivité de  $a$  nous assure que le dernier terme est toujours positif. On a donc :

$$J(w) = J(u) + \frac{1}{2}a(w_1, w_1) \geq J(u)$$

et donc  $J(u)$  est certainement inférieure ou égale à  $J(w)$ . Puisque ce raisonnement est valide quel que soit  $w \in V$ ,  $u$  minimise bien la fonctionnelle  $J$  sur l'espace  $V$ .

Inversement, si  $u$  minimise  $J$ , on démontre que  $u$  est aussi une solution du problème variationnel 3.4. Considérons pour ce faire la fonction de la variable réelle  $\epsilon$  définie par :

$$g(\epsilon) = J(u + \epsilon w) = \frac{1}{2}a(u + \epsilon w, u + \epsilon w) - l(u + \epsilon w)$$

Puisque  $u$  minimise  $J$ , la fonction  $g$  possède un minimum local en  $\epsilon = 0$ , et ce quel que soit  $w$ . On doit donc avoir  $g'(0) = 0$ , quel que soit  $w$ . Mais en développant, on trouve :

$$g(\epsilon) = J(u) + \epsilon (a(u, w) - l(w)) + \frac{\epsilon^2}{2}a(w, w)$$

et donc :

$$g'(\epsilon) = a(u, w) - l(w) + \epsilon a(w, w) \quad \text{et} \quad g'(0) = a(u, w) - l(w)$$

La condition  $g'(0) = 0$  quel que soit  $w$  est donc équivalente à l'équation 3.4. ★

### 3.3 Applications

Nous commençons à aborder dans cette section les applications du théorème de Lax-Milgram à la résolution d'équations aux dérivées partielles. Pour simplifier l'exposé, nous abordons séparément les équations aux dérivées partielles d'ordre 2 et celles d'ordre 4. Nous verrons en effet que les espaces de Sobolev impliqués sont respectivement les espaces  $H^1(\Omega)$  et  $H^2(\Omega)$  et certains de leurs sous-espaces.

#### 3.3.1 Problèmes d'ordre 2

Bon nombre d'applications importantes résultent en des équations aux dérivées partielles d'ordre 2. Nous allons établir dans cette section que le cadre fonctionnel approprié est l'espace  $H^1(\Omega)$  et ses variantes comme  $H_0^1(\Omega)$ . L'imposition des conditions aux limites dictera le choix précis de l'espace  $V$ .

##### ► Exemple 3.5

Nous avons maintenant tous les moyens nécessaires pour revenir sur le problème présenté au chapitre 1. Rappelons que l'équation différentielle est :

$$\begin{cases} -u''(x) = f(x) \\ u(0) = u(1) = 0 \end{cases} \quad (3.6)$$

et que nous avons obtenu la formulation variationnelle :

$$\int_0^1 u'(x)w'(x)dx = \int_0^1 f(x)w(x)dx \quad \forall w(x) \mid w(0) = w(1) = 0 \quad (3.7)$$

En regardant de plus près la formulation variationnelle, on constate que seule la dérivée d'ordre 1 de  $u$  (et de la fonction test  $w$ ) est présente. On en conclut que l'espace  $H^1(]0, 1[)$  a toutes les chances de répondre à nos besoins c.-à-d. de nous permettre de vérifier les hypothèses du théorème de Lax-Milgram. De plus, les conditions aux limites sur  $u$  nous forcent à choisir  $H_0^1(]0, 1[)$ . Le terme de droite sera défini si, par exemple, la fonction  $f(x)$  est dans  $L^2(]0, 1[)$  ou plus généralement si  $f \in H^{-1}(\Omega)$ . On reformule donc le problème sous la forme :

♠ pour  $f(x)$  dans  $L^2(\Omega)$ , trouver  $u$  dans  $H_0^1(]0, 1[)$  telle que :

$$\int_0^1 u'(x)w'(x)dx = \int_0^1 f(x)w(x)dx \quad \forall w \in H_0^1(]0, 1[) \quad (3.8)$$

Vérifions maintenant les hypothèses du théorème de Lax-Milgram. On pose donc  $V = H_0^1(]0, 1[)$  et :

$$a(u, w) = \int_0^1 u'(x)w'(x)dx \quad \text{et} \quad l(w) = \int_0^1 f(x)w(x)dx$$

Nous avons déjà établi la continuité de la forme linéaire  $l(w)$ . La bilinéarité de la forme  $a$  est triviale à démontrer. De plus, cette forme bilinéaire est symétrique et continue. En effet, en se servant une fois encore de l'inégalité de Cauchy, on a :

$$|a(u, w)| = \left| \int_0^1 u'(x)w'(x)dx \right| \leq \|u'\|_{0,\Omega} \|w'\|_{0,\Omega} = |u|_{1,\Omega} |w|_{1,\Omega}$$

et le résultat suit puisque  $|\cdot|_{1,\Omega}$  est une norme sur  $H_0^1(]0, 1[)$ . Enfin :

$$a(w, w) = \int_0^1 w'(x)w'(x)dx = |w|_{1,\Omega}^2$$

et la coercivité suit en posant  $\alpha = 1$ . ◀

### ◆ Remarque 3.2

Dans l'exemple précédent, nous avons utilisé la norme  $|\cdot|_{1,\Omega}$  pour vérifier les hypothèses du théorème de Lax-Milgram. Cela était rendu possible car  $|\cdot|_{1,\Omega}$  et  $\|\cdot\|_{1,\Omega}$  sont des normes équivalentes sur l'espace  $H_0^1(]0, 1[)$ . En pratique, sur un espace  $V$  quelconque, on vérifie les hypothèses du théorème de Lax-Milgram en utilisant parmi différentes normes équivalentes sur l'espace  $V$  celle qui nous facilite le plus le travail. ◆

### ► Exemple 3.6

Cet exemple est simplement une généralisation en plusieurs dimensions de l'exemple précédent. Soit donc  $f(\mathbf{x})$  une fonction de  $L^2(\Omega)$  et considérons l'équation aux dérivées partielles :

$$\begin{cases} -\nabla \cdot (k\nabla u) = f & \text{dans } \Omega \\ u = 0 & \text{sur } \Gamma \end{cases}$$

On peut aisément donner plusieurs interprétations physiques à ce problème. On peut par exemple considérer  $u$  comme une température (en °C),  $k$  comme étant la conductivité thermique (en W/(m °C)) que nous supposons constante (et strictement positive) et  $f(\mathbf{x})$  comme une source d'énergie (en W/m<sup>3</sup>). La température est maintenue à 0 sur la paroi du domaine  $\Omega$ .

Il paraît encore ici naturel de considérer l'espace  $H_0^1(\Omega)$  puisque les conditions aux limites sur  $u$  sont homogènes (nulles). En multipliant par une fonction test  $w \in H_0^1(\Omega)$ , on trouve à l'aide du théorème de la divergence A.5 :

$$\int_{\Omega} k\nabla u \cdot \nabla w \, dv - \int_{\Gamma} k \frac{\partial u}{\partial \mathbf{n}} w \, ds = \int_{\Omega} f w \, dv \quad \forall w \in H_0^1(\Omega)$$

L'intégrale de bord fait intervenir deux termes. Tout d'abord le terme  $k \frac{\partial u}{\partial \mathbf{n}}$  qui apparaît naturellement, suite à l'utilisation du théorème de la divergence. Pour cette raison, on dit que c'est une *condition aux limites naturelle* du problème. L'autre terme est la fonction test  $w$ . Puisque  $w \in H_0^1(\Omega)$ , l'intégrale de bord s'annule (mais ce ne sera pas toujours le cas) et on obtient la formulation variationnelle :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv = \int_{\Omega} f w \, dv \quad \forall w \in H_0^1(\Omega) \quad (3.9)$$

qui est une généralisation multidimensionnelle de l'exemple précédent. On vérifie alors les hypothèses du théorème de Lax-Milgram de manière similaire. ◀

### ◆ Remarque 3.3

Le problème variationnel 3.9 est équivalent au problème de minimisation :

$$J(u) = \min_{w \in H_0^1(\Omega)} J(w) = \min_{w \in H_0^1(\Omega)} \frac{1}{2} \int_{\Omega} k |\nabla w|^2 \, dv - \int_{\Omega} f w \, dv$$

puisque la forme bilinéaire :

$$a(u, w) = \int_{\Omega} k \nabla u \cdot \nabla w \, dv$$

est symétrique. Les unités de cette fonctionnelle sont des W °C. On utilise cependant rarement cette interprétation en pratique et on préfère travailler directement avec la formulation variationnelle 3.9.

◆

### ◆ Remarque 3.4

Dans cet exemple, nous avons choisi la fonction  $f(\mathbf{x})$  dans l'espace  $L^2(\Omega)$  pour simplifier l'exposé. En fait, ce problème est bien posé si on remplace la fonction  $f$  par une distribution  $T$  appartenant au dual de  $H_0^1(\Omega)$  c.-à-d.  $T \in H^{-1}(\Omega)$ . Il faudrait alors remplacer le terme de droite dans la formulation variationnelle par  $\langle T, w \rangle$  puisque rien ne nous permet de supposer que cette distribution peut s'écrire sous une forme intégrale. ◆

Les deux exemples précédents sont particuliers en ce sens que les conditions aux limites sont nulles dans les deux cas. Il est temps de voir ce qui arrive dans le cas de conditions aux limites non homogènes. On se servira ici de certains résultats établis au chapitre précédent.

### ► Exemple 3.7

Considérons le problème :

$$\begin{cases} -\nabla \cdot (k \nabla u) = f & \text{dans } \Omega \\ u = g & \text{sur } \Gamma \end{cases}$$

On peut interpréter physiquement ce problème comme pour l'exemple précédent à l'exception près que la température à la paroi est maintenue à une valeur  $g(\mathbf{x})$  éventuellement non nulle. Nous résoudrons ce problème de deux façons différentes qui aboutiront bien sûr à la même formulation variationnelle. Cela nous permettra de montrer comment on traite les conditions aux limites non homogènes.

**Première approche :**

Tout d'abord, il convient de noter que la fonction  $g$  se doit d'appartenir à l'espace  $H^{1/2}(\Gamma)$  pour que le problème ait un sens. De plus, il serait tentant de considérer :

$$V = \{w_1 \in H^1(\Omega) \mid w_1 = g \text{ sur } \Gamma\} \quad (3.10)$$

mais ce n'est pas un espace fonctionnel linéaire. La clé vient de l'équation 2.30 qui assure l'existence d'une fonction  $u_g \in H^1(\Omega)$  dont  $g$  est la trace au bord. En ce sens, on dit que pour un problème d'ordre 2, l'imposition de  $u$  sur la frontière (ou une partie de la frontière) est une condition aux limites *essentielle*. On appelle cette étape le *relèvement des conditions aux limites essentielles*. On décompose alors la fonction recherchée  $u$  en une somme  $\delta_u + u_g$  où  $\delta_u \in H_0^1(\Omega)$  et s'annule donc sur  $\Gamma$ . On considère ensuite le problème équivalent :

$$\begin{cases} -\nabla \cdot (k \nabla \delta_u) = f + \nabla \cdot (k \nabla u_g) & \text{dans } \Omega \\ \delta_u = 0 & \text{sur } \Gamma \end{cases}$$

On s'est donc ramené une fois de plus à un problème avec conditions aux limites homogènes. Il est facile de vérifier que les deux formulations fortes sont équivalentes. En multipliant par une fonction test  $w \in H_0^1(\Omega)$  et en intégrant par parties, on trouve, à l'aide du théorème de la divergence A.5 :

$$\int_{\Omega} k \nabla \delta_u \cdot \nabla w \, dv - \int_{\Gamma} k \frac{\partial \delta_u}{\partial \mathbf{n}} w \, ds = \int_{\Omega} f w \, dv - \int_{\Omega} k \nabla u_g \cdot \nabla w \, dv + \int_{\Gamma} k \frac{\partial u_g}{\partial \mathbf{n}} w \, ds$$

Puisque  $w \in H_0^1(\Omega)$ , les intégrales de bord s'annulent et on obtient la formulation variationnelle suivante :

♠ trouver  $\delta_u \in H_0^1(\Omega)$  telle que :

$$\int_{\Omega} k \nabla \delta_u \cdot \nabla w \, dv = \int_{\Omega} f w \, dv - \int_{\Omega} k \nabla u_g \cdot \nabla w \, dv \quad \forall w \in H_0^1(\Omega) \quad (3.11)$$

Ce problème variationnel requiert la connaissance de la fonction de relèvement des conditions aux limites  $u_g$ . Or jusqu'à maintenant, nous savons qu'une telle fonction existe mais sans plus. La méthode des éléments finis nous permettra de construire explicitement cette fonction. Nous pouvons donc procéder comme si cette fonction était connue. La vérification des hypothèses du théorème de Lax-Milgram ne pose aucune difficulté particulière. On pose :

$$a(\delta_u, w) = \int_{\Omega} k \nabla \delta_u \cdot \nabla w \, dv \quad \text{et} \quad l(w) = \int_{\Omega} f w \, dv - \int_{\Omega} k \nabla u_g \cdot \nabla w \, dv$$

La continuité, la bilinéarité et la coercivité de  $a$  sur  $H_0^1(\Omega)$  ont déjà été établies. La linéarité de  $l$  découle de la linéarité des intégrales. Quant à la continuité, on a :

$$\begin{aligned} |l(w)| &\leq \|f\|_{0,\Omega} \|w\|_{0,\Omega} + k \sum_{i=1}^3 \left\| \frac{\partial u_g}{\partial x_i} \right\|_{0,\Omega} \left\| \frac{\partial w}{\partial x_i} \right\|_{0,\Omega} \\ &\leq \|f\|_{0,\Omega} \|w\|_{1,\Omega} + k \sum_{i=1}^3 \left\| \frac{\partial u_g}{\partial x_i} \right\|_{0,\Omega} \|w\|_{1,\Omega} \end{aligned}$$

c.-à-d.

$$|l(w)| \leq C \|w\|_{1,\Omega} \quad \text{où } C = \|f\|_{0,\Omega} + k \sum_{i=1}^3 \left\| \frac{\partial u_g}{\partial x_i} \right\|_{0,\Omega}$$

### Deuxième approche :

On part tout simplement de l'équation différentielle initiale. Puisque qu'on impose une condition aux limites essentielle sur toute la frontière, on choisit des fonctions tests  $w$  s'annulant sur toute la frontière ce qui revient à choisir  $V = H_0^1(\ ]0, 1[ )$ . On multiplie ensuite par  $w \in V$  et on intègre par parties :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv - \int_{\Gamma} k \frac{\partial u}{\partial \mathbf{n}} w \, ds = \int_{\Omega} f w \, dv$$

Puisque  $w \in H_0^1(\Omega)$ , les intégrales de bord s'annulent et on obtient la formulation variationnelle suivante :

♠ trouver  $u \in \{w_1 \in H^1(\Omega) \mid w_1 = g \text{ sur } \Gamma\}$  telle que :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv = \int_{\Omega} f w \, dv \quad \forall w \in H_0^1(\Omega) \quad (3.12)$$

On ne peut appliquer immédiatement le théorème de Lax-Milgram puisque la solution  $u$  et les fonctions tests  $w$  ne sont pas dans le même espace (en fait  $u$  n'est même pas dans un espace fonctionnel linéaire). C'est ici que l'on peut effectuer le relèvement des conditions aux limites essentielles. On pose comme précédemment  $u = \delta_u + u_g$  où  $u_g$  vérifie les conditions aux limites essentielles et on obtient immédiatement la formulation variationnelle 3.11.

### Unicité de $u$ :

Nous avons donc démontré l'existence et l'unicité de  $\delta_u$  pour un relèvement  $u_g$  des conditions aux limites essentielles. La solution de notre problème initial est donc  $u = \delta_u + u_g$ . On remarque cependant que  $u_g$  n'est pas unique. Se peut-il que  $u$  ne soit pas unique ? La formulation 3.12 nous permet cependant de démontrer facilement l'unicité de  $u$ . En effet, si on suppose qu'il existe deux solutions  $u_1$  et  $u_2$  de l'équation 3.12, alors la différence  $u_1 - u_2$  est dans  $H_0^1(\Omega)$  car  $u_1 - u_2 = g - g = 0$  sur  $\Gamma$  et de plus :

$$\int_{\Omega} k \nabla (u_1 - u_2) \cdot \nabla w \, dv = 0 \quad \forall w \in H_0^1(\Omega)$$

En choisissant  $w = (u_1 - u_2)$ , et puisque  $k > 0$ , on montre que  $|u_1 - u_2|_{1,\Omega} = 0$  ce qui entraîne que  $u_1 \stackrel{\text{p.p.}}{=} u_2$  puisque  $|\cdot|_{1,\Omega}$  est une norme. ◀

### ◆ Remarque 3.5

Pour les équations aux dérivées partielles d'ordre 2, l'imposition de  $u$  sur la frontière (ou une partie de la frontière) constitue une condition aux limites essentielle qui peut être relevée en vertu de l'équation 2.30. Là où la condition essentielle est imposée, la fonction test  $w$  doit s'annuler. Les fonctions tests  $w$  sont parfois appelées des *déplacements admissibles*, en référence à l'origine de la méthode des éléments finis en mécanique des structures. ◆

### ► Exemple 3.8

On considère cette fois un problème dit *mêlé*, en ce sens que les conditions aux limites sont de deux types. Soit donc le problème :

$$\left\{ \begin{array}{l} -\nabla \cdot (k\nabla u) = f \quad \text{dans } \Omega \\ u = g \quad \text{sur } \Gamma_0 \quad (g \in H^{1/2}(\Gamma_0)) \\ k \frac{\partial u}{\partial \mathbf{n}} = h \quad \text{sur } \Gamma_1 \end{array} \right.$$

où les frontières  $\Gamma_0$  et  $\Gamma_1$  vérifient  $\Gamma_0 \cap \Gamma_1 = \emptyset$  et  $\Gamma_0 \cup \Gamma_1 = \Gamma$ . La fonction  $h$  s'interprète comme un flux de chaleur à travers la frontière (en  $\text{W/m}^2$ ).

Le relèvement de la condition aux limites essentielle est encore possible. En vertu de l'équation 2.30, on peut trouver une fonction  $u_g \in H^1(\Omega)$  telle que  $u_g = g$  sur  $\Gamma_0$ . Nous ne pouvons pas faire le même exercice pour la condition aux limites de type Neumann sur la dérivée normale. Nous n'avons en effet aucune assurance de l'existence d'une fonction  $u_h \in H^1(\Omega)$  telle que :

$$k \frac{\partial u_h}{\partial \mathbf{n}} = h \quad \text{sur } \Gamma_1$$

Puisqu'on impose une condition essentielle uniquement sur  $\Gamma_0$ , on choisit  $V = H_{\Gamma_0}^1(\Omega)$ . On multiplie ensuite par une fonction test  $w \in V$  et on intègre en utilisant le théorème de la divergence :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv - \int_{\Gamma} k \frac{\partial u}{\partial \mathbf{n}} w \, ds = \int_{\Omega} f w \, dv$$

Puisque  $w \in H_{\Gamma_0}^1(\Omega)$ , l'intégrale de bord s'annulent sur  $\Gamma_0$  tandis que sur  $\Gamma_1$ , on a :

$$k \frac{\partial u}{\partial \mathbf{n}} = h$$

qui est la condition naturelle. Il reste, après substitution :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv = \int_{\Omega} f w \, dv + \int_{\Gamma_1} h w \, ds$$

On pose alors comme précédemment  $u = \delta_u + u_g$  où  $\delta_u \in H_{\Gamma_0}^1(\Omega)$  et  $u_g = g$  sur  $\Gamma_0$ . On a alors le problème équivalent :

$$\int_{\Omega} k \nabla \delta_u \cdot \nabla w \, dv = \int_{\Omega} f w \, dv - \int_{\Omega} k \nabla u_g \cdot \nabla w \, dv + \int_{\Gamma_1} h w \, ds \quad (3.13)$$

qui est bien de la forme du théorème de Lax-Milgram.

♠ trouver  $\delta_u \in H_{\Gamma_0}^1(\Omega)$  telle que :

$$a(\delta_u, w) = l(w) \quad \forall w \in H_{\Gamma_0}^1(\Omega)$$

L'espace  $V$  est ici  $H_{\Gamma_0}^1(\Omega)$  car les conditions aux limites essentielles l'imposent. Rappelons encore ici, car cela est important, que la méthode des éléments finis nous permettra de construire explicitement la fonction  $u_g$ . La vérification des hypothèses du théorème de Lax-Milgram est laissée au lecteur sauf pour un point qui mérite une attention particulière. Pour démontrer la continuité de  $l(w)$ , on doit recourir à la continuité de la trace au bord et à l'équation 2.31 en effectuant la majoration :

$$\left| \int_{\Gamma_1} h w \, ds \right| \leq \|h\|_{0,\Gamma_1} \|w\|_{0,\Gamma_1} \leq C \|h\|_{0,\Gamma_1} \|w\|_{1,\Omega}$$

◀

### ◆ Remarque 3.6

Dans le terme de bord :

$$\int_{\Gamma} k \frac{\partial u}{\partial \mathbf{n}} w \, ds$$

on doit connaître soit la condition naturelle, dans ce cas  $k \frac{\partial u}{\partial \mathbf{n}}$ , soit la condition essentielle, dans ce cas  $u$  et alors la fonction test  $w$  s'annule. On ne peut pas imposer au même endroit de la frontière la condition essentielle et la condition naturelle. On peut faire l'analogie avec une force et un déplacement. On connaît soit la force (condition naturelle) soit le déplacement (condition essentielle). ◆

### 3.3.2 Problèmes d'ordre 4

Nous considérons ici les équations aux dérivées partielles d'ordre 4 de la forme :

$$\frac{d^2}{dx^2} \left( q(x) \frac{d^2 u}{dx^2} \right) = f(x)$$

ou encore en dimension 2 ou 3 :

$$\nabla^2 (q(\mathbf{x}) \nabla^2 u(\mathbf{x})) = f(\mathbf{x})$$

auxquelles s'ajoutent des conditions aux limites appropriées. Il est facile de se convaincre qu'après deux utilisations du théorème de la divergence (deux intégrations par parties en dimension 1), il ne restera dans la formulation variationnelle que des dérivées d'ordre 2. L'espace fonctionnel approprié a donc toutes les chances d'être  $H^2(\Omega)$  ou ses variantes. Voyons tout cela de plus près en commençant un problème en dimension 1.

► **Exemple 3.9**

$$\left\{ \begin{array}{l} \frac{d^2}{dx^2} \left( q(x) \frac{d^2 u}{dx^2} \right) = f(x) \text{ dans } ]0, L[ \\ u(0) = \frac{du}{dx}(0) = 0 \\ q(x) \frac{d^2 u}{dx^2} \Big|_{x=L} = M_0 \\ \frac{d}{dx} \left( q(x) \frac{d^2 u}{dx^2} \right) \Big|_{x=L} = 0 \end{array} \right.$$

Ce problème provient de l'analyse de la déformation d'une poutre sous l'effet d'une sollicitation  $f(x)$  (en N/m) et d'un moment de flexion  $M_0$  (en N · m). La fonction  $q(x)$  dépend alors des propriétés élastiques du matériau et de l'aire de la section de la poutre. En fait on a  $q(x) = EI$  où  $E$  est le module d'Young et  $I$  le moment d'inertie. Cette fonction (pas forcément constante le long de la poutre) est en général strictement positive et bornée de sorte qu'il existe des constantes  $q_1$  et  $q_2$  telles que :

$$0 < q_1 \leq q(x) \leq q_2 \quad \forall x \in [0, L]$$

Pour un problème d'ordre 4, le théorème 2.9 nous assure que les conditions aux limites essentielles portent sur  $u$  et  $\frac{du}{dx}$  car on peut en effectuer le relèvement. Notez la différence importante avec les problèmes d'ordre 2 pour lesquels la seule condition aux limites essentielle est sur  $u$ . Dans cet exemple, la fonction de relèvement est tout simplement  $u_g = 0$  puisque les conditions essentielles sont homogènes. L'espace fonctionnel approprié est donc :

$$V = \left\{ w \in H^2(]0, L[) \mid w(0) = \frac{dw}{dx}(0) = 0 \right\}$$

On multiplie alors par une fonction test  $w \in V$  et on intègre une première fois par parties. On obtient ainsi :

$$-\int_0^L \frac{d}{dx} \left( q(x) \frac{d^2 u}{dx^2} \right) \frac{dw}{dx} dx + \left( \frac{d}{dx} \left( q(x) \frac{d^2 u}{dx^2} \right) \right) w \Big|_0^L = \int_0^L f w dx$$

Le terme de bord fait intervenir d'une part la fonction test  $w$  qui s'annule en  $x = 0$  et d'autre part la première condition naturelle :

$$\frac{d}{dx} \left( q(x) \frac{d^2 u}{dx^2} \right) \quad (3.14)$$

qui vaut 0 en  $x = L$ . Il reste :

$$- \int_0^L \frac{d}{dx} \left( q(x) \frac{d^2 u}{dx^2} \right) \frac{dw}{dx} dx = \int_0^L f w dx$$

Une nouvelle intégration par parties résulte en :

$$\int_0^L q(x) \frac{d^2 u}{dx^2} \frac{d^2 w}{dx^2} dx - \left( q(x) \frac{d^2 u}{dx^2} \right) \frac{dw}{dx} \Big|_0^L = \int_0^L f w dx$$

Le nouveau terme de bord fait intervenir  $\frac{dw}{dx}$  qui s'annulent en  $x = 0$  et la deuxième condition dite naturelle :

$$q(x) \frac{d^2 u}{dx^2} \quad (3.15)$$

qui prend la valeur  $M_0$  en  $x = L$ . On a donc la formulation variationnelle :

$$\int_0^L q(x) \frac{d^2 u}{dx^2} \frac{d^2 w}{dx^2} dx = \int_0^L f w dx + M_0 \frac{dw}{dx}(L) \quad (3.16)$$

qui est de la forme  $a(u, w) = l(w)$ .

Vérifions les hypothèses du théorème de Lax-Milgram. La linéarité de  $l$  et la bilinéarité de  $a$  sont évidentes de même que sa symétrie. Pour démontrer la continuité de  $a$ , on utilise le fait que la fonction  $q(x)$  est bornée et comme toujours, l'inégalité de Cauchy. On a :

$$|a(u, w)| = \left| \int_0^L q(x) \frac{d^2 u}{dx^2} \frac{d^2 w}{dx^2} dx \right| \leq q_2 \int_0^L \left| \frac{d^2 u}{dx^2} \frac{d^2 w}{dx^2} \right| dx \leq q_2 \|u''\|_{0,\Omega} \|w''\|_{0,\Omega}$$

et le résultat suit puisque  $\|w''\|_{0,\Omega}$  est une norme sur  $V$  (voir l'équation 2.38). En ce qui concerne la coercivité, le raisonnement est similaire. On a :

$$a(w, w) = \int_0^L q(x) \frac{d^2 w}{dx^2} \frac{d^2 w}{dx^2} dx \geq q_1 \int_0^L \frac{d^2 w}{dx^2} \frac{d^2 w}{dx^2} dx = q_1 \|w''\|_{0,\Omega}^2 = q_1 |w|_{2,\Omega}^2$$

et le résultat suit en prenant  $\alpha = q_1$  dans la définition 3.3 et puisque  $|w|_{2,\Omega}^2$  est une norme équivalente à la norme  $\|w\|_{2,\Omega}^2$  sur  $V$ .

Enfin, il reste à montrer la continuité de la forme linéaire  $l$ . Pour y arriver, on doit se servir de l'inégalité de Cauchy et de la continuité de la trace au bord 2.36. On a ainsi :

$$\begin{aligned} |l(w)| &= \left| \int_0^L f w dx + M_0 \frac{dw}{dx}(L) \right| \leq \left| \int_0^L f w dx \right| + |M_0| \left| \frac{dw}{dx}(L) \right| \\ &\leq \|f\|_{0,\Omega} \|w\|_{0,\Omega} + |M_0| \|w\|_{2,\Omega} \leq (\|f\|_{0,\Omega} + |M_0|) \|w\|_{2,\Omega} \end{aligned}$$



◆ **Remarque 3.7**

Le problème variationnel 3.16 est équivalent à la minimisation de la fonctionnelle :

$$J(u) = \min_{w \in V} J(w) = \min_{w \in V} \frac{1}{2} \int_0^L EI \left( \frac{d^2 w}{dx^2} \right)^2 dx - \int_0^L f w dx - M_0 \frac{dw}{dx}(L)$$

dont les unités sont des N·m. Le premier terme de  $J(w)$  correspond à l'énergie élastique de déformation tandis que les deux autres termes correspondent au travail effectué par la charge imposée.



◆ **Remarque 3.8**

Pour les problèmes d'ordre 4, il y a 2 conditions aux limites essentielles  $u$  et  $\frac{du}{dx}$  ainsi que 2 conditions aux limites naturelles qui sont en relation étroite. Dans le premier terme de bord :

$$\left( \frac{d}{dx} \left( q(x) \frac{d^2 u}{dx^2} \right) \right) w \Big|_0^L$$

on doit connaître soit la condition naturelle  $\frac{d}{dx} \left( q(x) \frac{d^2 u}{dx^2} \right)$  auquel cas la fonction test  $w$  ne doit pas s'annuler, ou encore on connaît  $u$  et alors la fonction test  $w$  s'annule à cet endroit. Un raisonnement similaire existe pour la deuxième condition naturelle. Si  $q(x) \frac{d^2 u}{dx^2}$  est donnée, alors  $\frac{dw}{dx}$  ne peut s'annuler. Par contre, si  $\frac{du}{dx}$  est connue, alors la fonction test  $\frac{dw}{dx}$  s'annule à cet endroit et le terme de bord correspondant disparaît. ◆

### 3.3.3 Résumé

On peut établir en quelque sorte la procédure à suivre pour vérifier les hypothèses du théorème de Lax-Milgram.

– **Problèmes d'ordre 2**

1. la seule condition essentielle est l'imposition de  $u$  ;
2. on identifie la portion de la frontière  $\Gamma_0$  où la condition essentielle est imposée ;
3. on pose  $V = H_{\Gamma_0}^1(\Omega)$ . Si aucune condition essentielle n'est imposée, on pose  $V = H^1(\Omega)$  ;
4. l'inclusion 2.29 nous assure de l'existence d'une fonction de relèvement  $u_g$  des conditions aux limites essentielles ;
5. là où la condition essentielle est imposée, les fonctions tests  $w$  s'annulent (par définition de l'espace  $V$ ) ;

6. on obtient la formulation variationnelle en multipliant par une fonction test  $w$  et en intégrant par parties les termes d'ordre 2 ;
7. la condition aux limites naturelle est le coefficient (dépendant de  $u$ ) qui multiplie  $w$  dans le terme de bord ;
8. on vérifie les hypothèses du théorème de Lax-Milgram en utilisant l'une ou l'autre des normes équivalentes sur l'espace  $V$ . La relation 2.31 sera éventuellement utile pour traiter les termes de bords ;
9. on en déduit (s'il y a lieu) l'existence et l'unicité de la solution  $u$ .

– **Problèmes d'ordre 4**

1. il y a 2 conditions essentielles soit l'imposition de  $u$  et de  $\partial u / \partial \mathbf{n}$  ( $u'(x)$  en dimension 1) ;
2. on identifie la portion de la frontière  $\Gamma_0$  où les conditions essentielles sont imposées ;
3. on pose  $V = H_{\Gamma_0}^2(\Omega)$ . Si aucune condition essentielle n'est imposée, on pose  $V = H^2(\Omega)$  ;
4. l'équation 2.35 nous assure de l'existence d'une fonction de relèvement  $u_g$  des conditions aux limites essentielles ;
5. là où la condition essentielle sur  $u(x)$  est imposée, les fonctions tests  $w$  s'annulent. Là où la condition essentielle sur  $\partial u / \partial \mathbf{n}$  est imposée, les dérivées normales des fonctions tests  $\partial w / \partial \mathbf{n}$  s'annulent (par définition de l'espace  $V$ ) ;
6. on obtient la formulation variationnelle en multipliant par une fonction test  $w$  et en intégrant par parties les termes d'ordre 4 (et parfois aussi des termes d'ordre 2) ;
7. les conditions aux limites naturelles sont les coefficients (dépendant de  $u$ ) qui multiplient  $w$  et  $\partial w / \partial \mathbf{n}$  dans les termes de bord ;
8. on vérifie les hypothèses du théorème de Lax-Milgram en utilisant l'une ou l'autre des normes équivalentes sur l'espace  $V$ . La relation 2.36 sera éventuellement utile pour traiter les termes de bords ;
9. on en déduit (s'il y a lieu) l'existence et l'unicité de la solution  $u$ .

### 3.4 Exercices

1. Vérifier si les expressions suivantes sont des formes bilinéaires continues, symétriques et coercives sur les espaces donnés. On supposera que les fonctions  $p(\mathbf{x})$  et  $q(\mathbf{x})$  sont bornées c'est-à-dire  $0 < p_1 \leq p(\mathbf{x}) \leq p_2$  et  $0 < q_1 \leq q(\mathbf{x}) \leq q_2$  et ce  $\forall \mathbf{x} \in \Omega$ .

a)

$$a(u, w) = \int_{\Omega} p(\mathbf{x})uw \, dv \text{ dans } L^2(\Omega)$$

b)

$$a(u, w) = \int_{\Omega} q(\mathbf{x})\nabla u \cdot \nabla w \, dv \text{ dans } H_0^1(\Omega)$$

c)

$$a(u, w) = \int_{\Omega} q(\mathbf{x})\nabla u \cdot \nabla w \, dv \text{ dans } H^1(\Omega)$$

d)

$$a(u, w) = \int_{\Omega} (p(\mathbf{x})uw + q(\mathbf{x})\nabla u \cdot \nabla w) \, dv \text{ dans } H^1(\Omega)$$

2. Discuter de l'existence et de l'unicité de la solution des problèmes suivants. Bien identifier l'espace fonctionnel  $V$ , les conditions aux limites essentielles et naturelles et obtenir explicitement la fonction de relèvement des conditions aux limites essentielles lorsque possible. La fonction  $q(\mathbf{x})$  sera supposée continue et bornée c.-à-d.  $0 < q_1 \leq q(\mathbf{x}) \leq q_2$  et  $\nu_1$  et  $\nu_2$  seront des constantes strictement positives. La fonction  $f(\mathbf{x})$  apparaissant à plusieurs reprises sera supposée dans  $L^2(\Omega)$ .

En dimension 1 :

a)

$$\begin{cases} -\nu_2 \frac{d^2 u}{dx^2} = f & \text{dans } ]0, 1[ \\ u(0) = u(1) = 0 \end{cases}$$

b)

$$\begin{cases} \nu_1 u - \nu_2 \frac{d^2 u}{dx^2} = f & \text{dans } ]0, 1[ \\ \frac{du}{dx}(0) = \frac{du}{dx}(1) = 0 \end{cases}$$

c)

$$\begin{cases} -\nu_2 \frac{d^2 u}{dx^2} = f & \text{dans } ]0, 1[ \\ u(0) = a, \quad u(1) = b \end{cases}$$

d)

$$\begin{cases} -\frac{d}{dx} \left( q(x) \frac{du}{dx} \right) = f & \text{dans } ]0, 1[ \\ u(0) = a, \quad q(1) \frac{du}{dx}(1) = b \end{cases}$$

e)

$$\begin{cases} \frac{d^2}{dx^2} \left( q(x) \frac{d^2u}{dx^2} \right) = f & \text{dans } ]0, L[ \\ u(0) = a, \quad u(L) = b \\ q(0) \frac{d^2u}{dx^2}(0) = c, \quad q(L) \frac{d^2u}{dx^2}(L) = d \end{cases}$$

En dimension supérieure à 1 :

f)

$$\begin{cases} -\nabla \cdot (q(\mathbf{x}) \nabla u) = f & \text{dans } \Omega \\ u = g & \text{sur } \Gamma_0 \\ q(\mathbf{x}) \frac{\partial u}{\partial \mathbf{n}} = h & \text{sur } \Gamma_1 \end{cases}$$

g)

$$\begin{cases} -\nabla \cdot (\nabla u) = f & \text{dans } \Omega \\ \frac{\partial u}{\partial \mathbf{n}} = h & \text{sur } \Gamma \end{cases}$$

h)

$$\begin{cases} -\nabla \cdot (\nabla u) = f & \text{dans } \Omega \\ u + \frac{\partial u}{\partial \mathbf{n}} = h & \text{sur } \Gamma \end{cases}$$

3. On souhaite résoudre le problème suivant :

$$\frac{d^2}{dx^2} \left( EI \frac{d^2u(x)}{dx^2} \right) = f(x) \text{ dans } ]0, 1[,$$

où  $EI$  est une constante. Peut-on obtenir une formulation variationnelle dans le cas des conditions aux limites suivantes :

$$u(0) = 0, \quad u'(0) = 0, \quad u(1) = 1, \quad \frac{d}{dx} \left( EI \frac{d^2u(1)}{dx^2} \right) = F$$

4. On suppose que l'on a un problème variationnel de la forme :

$$a(u, w) = l(w)$$

où la forme bilinéaire  $a$  et la forme linéaire  $l$  vérifient les hypothèses du théorème de Lax-Milgram dans un espace de Hilbert  $V$ .

Montrer que l'unique solution  $u$  dépend continûment des données c.-à-d. que :

$$\|u\|_V \leq \frac{\|l\|}{\alpha}$$

où  $\alpha$  est la constante de coercivité et  $\|l\|$  est la plus petite constante  $C$  telle que :

$$|l(w)| \leq C\|w\|_V \quad \forall w \in V \quad \text{ou encore} \quad |l(w)| \leq \|l\| \|w\|_V \quad \forall w \in V$$

5. Vérifier les hypothèses du théorème de Lax-Milgram pour le problème suivant :

$$\begin{cases} -\frac{d}{dx} \left( c_1 \frac{du}{dx} \right) + c_2 \frac{du}{dx} = f \\ u(0) = u(1) = 0 \end{cases}$$

où  $c_1$  et  $c_2$  sont des constantes strictement positives et  $f(x)$  est une fonction de  $L^2(]0, 1[)$ . Déterminer l'espace fonctionnel approprié et obtenir la formulation variationnelle du problème.

**Ne pas intégrer par parties le terme  $c_2 \frac{du}{dx}$ .**

**N.B.** : Passer rapidement sur les questions de linéarité et bilinéarité. On se rappellera de plus que :

$$g(x)g'(x) = \frac{1}{2} \frac{d}{dx} (g^2)$$



# Chapitre 4

## Méthode de Ritz

### 4.1 Principes généraux

Une fois le problème formulé sous forme variationnelle, il reste à le discrétiser c.-à-d. à le faire passer d'un problème de dimension infinie à un problème approché de dimension finie. Ce problème discrétisé sera ensuite résolu par les techniques d'algèbre linéaire classiques : résolution de systèmes algébriques linéaires ou non linéaires, de problèmes aux valeurs propres, etc.

La méthode de Ritz est une technique de discrétisation de problèmes variationnels et est en quelque sorte le précurseur de la méthode des éléments finis. Soit donc un problème variationnel vérifiant les hypothèses du théorème de Lax-Milgram :

♠ trouver  $u \in V$  telle que :

$$a(u, w) = l(w) \quad \forall w \in V \quad (4.1)$$

où la fonction  $u$  vérifie les conditions aux limites essentielles homogènes (le cas des conditions aux limites non homogènes nécessite la construction d'une fonction de relèvement  $u_g$  mais ne pose pas de difficultés théoriques supplémentaires). On se donne maintenant  $N$  fonctions  $\phi_j \in V, j = 1, 2, \dots, N$  appelées *fonctions d'interpolation de Ritz* ou plus simplement *fonctions de Ritz* vérifiant elles aussi les conditions essentielles homogènes. On suppose ensuite que l'on peut écrire :

$$u(\mathbf{x}) \simeq u^N(\mathbf{x}) = \sum_{j=1}^N u_j \phi_j(\mathbf{x}) \quad (4.2)$$

Dans cette expression, les  $N$  coefficients  $u_j$  sont à déterminer et le problème est maintenant de dimension finie  $N$ . L'ensemble de toutes les combinaisons linéaires possibles des fonctions  $\phi_j$  forme un sous-espace de dimension  $N$  de  $V$  noté  $V^N$  (toujours en supposant que les fonctions  $\phi_j$  sont choisies dans  $V$  dès le départ et qu'elles sont linéairement indépendantes). On considère donc l'approximation suivante du problème variationnel 4.1 :

♠ trouver  $u^N \in V^N$  telle que :

$$a(u^N, w^N) = l(w^N) \quad \forall w^N \in V^N \quad (4.3)$$

ou encore :

$$a \left( \sum_{j=1}^N u_j \phi_j, w^N \right) = l(w^N) \quad \forall w^N \in V^N$$

La bilinéarité de  $a$  nous permet d'écrire :

$$\sum_{j=1}^N u_j a(\phi_j, w^N) = l(w^N) \quad \forall w^N \in V^N$$

On va maintenant construire un système linéaire (parce que le problème de départ 4.1 est linéaire) dont les inconnues sont les coefficients  $u_j$ . Soit donc  $N$  nouvelles fonctions  $\tilde{\phi}_i, i = 1, 2, \dots, N$  appartenant à l'espace  $V^N$ . Puisque l'équation précédente est vraie quelle que soit la fonction  $w^N \in V^N$ , elle est valide pour chacune des fonctions  $\tilde{\phi}_i$  et on a :

$$\sum_{j=1}^N u_j a(\phi_j, \tilde{\phi}_i) = l(\tilde{\phi}_i) \quad 1 \leq i \leq N$$

qui est un système linéaire  $N$  sur  $N$  de la forme :

$$AU = F$$

où la matrice  $A$  et les vecteurs  $U$  et  $F$  ont pour coefficients :

$$\begin{bmatrix} a(\phi_1, \tilde{\phi}_1) & a(\phi_2, \tilde{\phi}_1) & \cdots & a(\phi_N, \tilde{\phi}_1) \\ a(\phi_1, \tilde{\phi}_2) & a(\phi_2, \tilde{\phi}_2) & \cdots & a(\phi_N, \tilde{\phi}_2) \\ \vdots & \vdots & \ddots & \vdots \\ a(\phi_1, \tilde{\phi}_N) & a(\phi_2, \tilde{\phi}_N) & \cdots & a(\phi_N, \tilde{\phi}_N) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} = \begin{bmatrix} l(\tilde{\phi}_1) \\ l(\tilde{\phi}_2) \\ \vdots \\ l(\tilde{\phi}_N) \end{bmatrix} \quad (4.4)$$

Si les fonctions  $\phi_j$  et  $\tilde{\phi}_i$  sont bien choisies, ce système linéaire est inversible et on peut dès lors déterminer les inconnues  $u_i$  et ainsi obtenir une approximation  $u^N(x)$  de la fonction  $u$  par la relation 4.2. Un choix naturel pour les fonctions  $\tilde{\phi}_i$  consiste à prendre tout simplement  $\tilde{\phi}_i = \phi_i$ , et ce pour tout  $i$ . C'est la *méthode de Ritz* ou *méthode de Rayleigh-Ritz*. Dans le cas où  $\tilde{\phi}_i \neq \phi_i$ , on parle de la *méthode de Petrov-Galerkin*.

Dans cet ouvrage, nous insisterons davantage sur la méthode de Ritz, bien qu'il existe des applications intéressantes de la méthode de Petrov-Galerkin, notamment pour les problèmes de convection-diffusion.

#### ◆ Remarque 4.1

On remarque que le coefficient  $A_{ij}$  du système linéaire 4.4 est de la forme :

$$A_{ij} = a(\phi_j, \tilde{\phi}_i)$$

Lorsque la forme bilinéaire est symétrique, il n'y a pas de confusion possible mais dans le cas général, il faut faire attention à l'ordre des indices  $i$  et  $j$ . ◆

## 4.2 Exemples

Nous voyons dans cette section quelques exemples d'applications de la méthode de Ritz. Nous en profiterons au passage pour mettre en évidence les forces et éventuellement les faiblesses de cette approche.

### ► Exemple 4.1

Considérons le problème en dimension 1 :

$$\begin{cases} -u''(x) &= \ln(x) \text{ dans } ]0, 1[ \\ u(0) &= 0 \\ \frac{du}{dx}(1) &= 1 \end{cases}$$

Puisque la seule condition essentielle (sur  $u$ ) est nulle, le relèvement est inutile ( $u_g = 0$ ) et on peut travailler directement avec  $u$  qui, dans ce cas, sera égal à  $\delta_u$ . La formulation variationnelle correspondante (laissée en exercice) est :

♠ trouver  $u \in V = \{w \in H^1(]0, 1[ \mid w(0) = 0\}$

$$\int_0^1 u'(x)w'(x)dx = w(1) + \int_0^1 \ln(x)w(x)dx \quad \forall w \in V \quad (4.5)$$

Les fonctions  $\phi_j$  doivent appartenir à l'espace  $V$  et vérifier les conditions aux limites essentielles homogènes. Dans ce cas, il suffit d'avoir  $\phi_j(0) = 0$ . Cela mis à part, ce choix est arbitraire si on s'assure que  $\phi_j \in V$ . On peut prendre par exemple :

$$\phi_j(x) = x^j$$

et on s'assure facilement que toutes les conditions sont bien remplies. Le coefficient général de la matrice  $A$  pour la méthode de Ritz est donc :

$$a_{ij} = \int_0^1 \phi_j'(x)\phi_i'(x) dx = \int_0^1 ijx^{i+j-2} dx = \frac{ij}{i+j-1}$$

tandis que le vecteur  $F$  a pour coefficients :

$$f_i = \phi_i(1) + \int_0^1 \ln(x)\phi_i(x) dx = 1 + \int_0^1 \ln(x)x^i dx = 1 - \frac{1}{(i+1)^2}$$

Ce système linéaire est ensuite résolu par les techniques habituelles de décomposition  $LU$  (voir par exemple Fortin, réf. [17]). En faisant varier la taille  $N$ , si tout se passe bien, on doit se rapprocher d'une éventuelle solution analytique. Dans cet exemple, cette solution est :

$$u(x) = \frac{3x^2}{4} - \frac{1}{2}x^2 \ln x$$

dont la dérivée est :

$$u'(x) = x - x \ln x$$

et on s'assure facilement que l'équation différentielle ainsi que les conditions aux limites sont vérifiées. En se servant du logiciel Matlab [25], on a résolu ce système pour obtenir les résultats de la figure 4.1 pour  $N = 1, 3$  et  $5$ . On y a superposé la solution analytique  $u(x)$  (en trait plein) et la solution numérique  $u^N(x)$ . Il est aussi intéressant de regarder le comportement de  $u'(x)$  et  $(u^N)'(x)$ . On remarque immédiatement qu'il est plus difficile d'approcher  $u'(x)$ . On peut affirmer que ce comportement est assez général. Nous en reparlerons lorsque nous aborderons les questions de convergence.

Enfin, il ne faut pas se leurrer sur la taille du système linéaire nécessaire pour obtenir une bonne approximation de  $u(x)$ . Dans cet exemple, une dimension de 5 semble suffire mais ce n'est certes pas toujours aussi facile... ◀

### ► Exemple 4.2

Considérons le problème en dimension 1 :

$$\begin{cases} -u''(x) + u & = 0 \text{ dans } ]0, 1[ \\ u(0) & = 1 \\ \frac{du}{dx}(1) + u(1) & = 0 \end{cases}$$

La condition essentielle n'est imposée qu'en  $x = 0$  et on choisit donc :

$$V = \{w \in H^1(]0, 1[) \mid w(0) = 0\}$$

On multiplie par  $w \in V$  et on intègre :

$$\int_0^1 (u'(x)w'(x) + u(x)w(x)) dx - \frac{du}{dx}w \Big|_0^1 = 0 \quad \forall w \in V$$

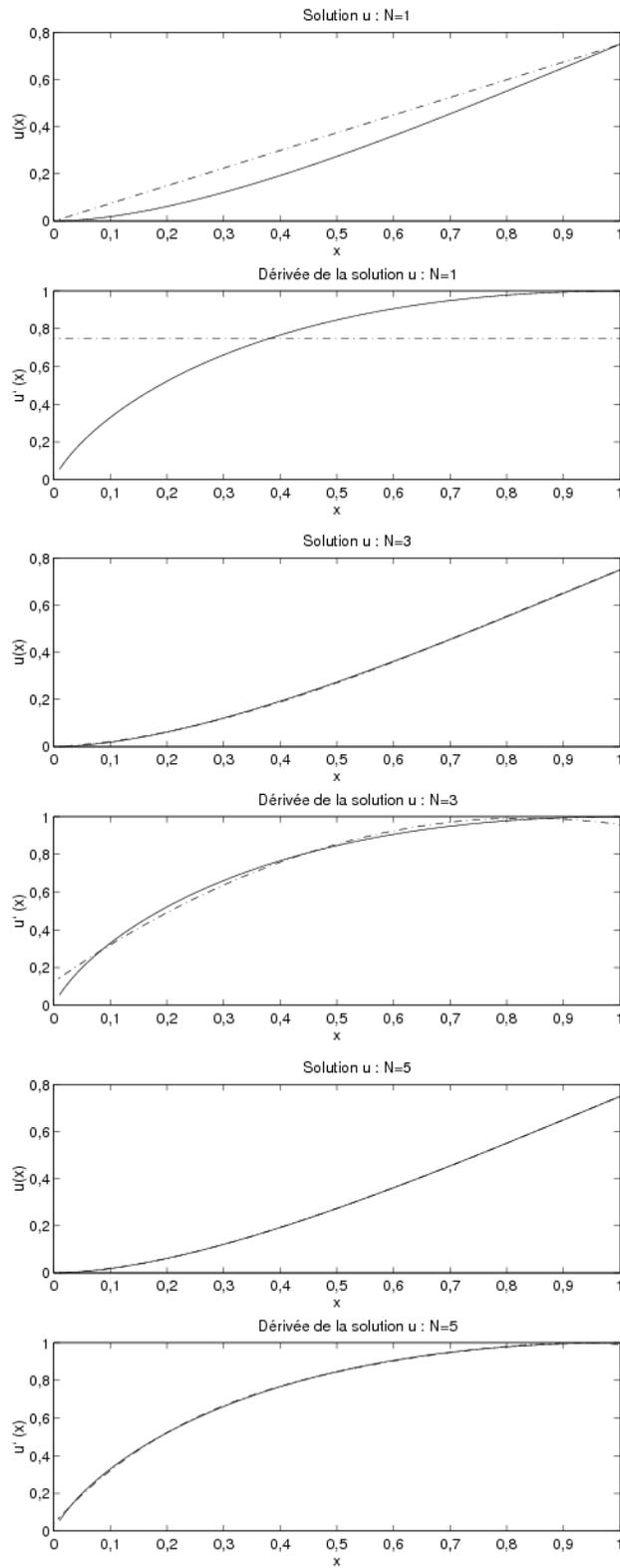
Le terme de bord est nul en  $x = 0$  mais en  $x = 1$ , on a  $\frac{du}{dx}(1) = -u(1)$ . Il en résulte :

$$\int_0^1 (u'(x)w'(x) + u(x)w(x)) dx + u(1)w(1) = 0 \quad \forall w \in V$$

On relève maintenant les conditions aux limites essentielles. On pose  $u = \delta_u + u_g$  et on peut choisir tout simplement la fonction  $u_g = 1$  qui appartient à  $H^1(]0, 1[)$  et qui satisfait la condition essentielle.

On a alors  $\forall w \in V$  :

$$\begin{aligned} & \int_0^1 (\delta'_u(x)w'(x) + \delta_u(x)w(x)) dx + \delta_u(1)w(1) \\ & = - \int_0^1 (u'_g(x)w'(x) + u_g(x)w(x)) dx - u_g(1)w(1) \end{aligned}$$

Figure 4.1 – Solutions  $u(x)$  et  $u'(x)$  pour  $N = 1$ ,  $N = 3$  et  $N = 5$

Puisque  $u_g = 1$ ,  $u'_g = 0$ , il reste :

$$\begin{aligned} & \int_0^1 (\delta'_u(x)w'(x) + \delta_u(x)w(x)) dx + \delta_u(1)w(1) \\ &= - \int_0^1 w(x)dx - w(1) \quad \forall w \in V \end{aligned}$$

qui est la formulation variationnelle correspondant au problème initial.

Pour la méthode de Ritz, on peut ici encore choisir  $\phi_i(x) = x^i, i = 1, 2, \dots, N$ . La matrice  $A$  correspondante a pour coefficients :

$$\begin{aligned} a_{ij} &= \int_0^1 (\phi'_j(x)\phi'_i(x) + \phi_j(x)\phi_i(x)) dx + \phi_j(1)\phi_i(1) \\ &= \int_0^1 (ijx^{i+j-2} + x^{i+j}) dx + 1 \\ &= \frac{ij}{i+j-1} + \frac{1}{i+j+1} + 1 \end{aligned}$$

tandis que le vecteur  $F$  a pour coefficients :

$$f_i = - \int_0^1 \phi_i(x) dx - \phi_i(1) = - \int_0^1 x^i dx - 1 = - \left( 1 + \frac{1}{(i+1)} \right)$$

Il est facile de s'assurer que la solution analytique de l'équation différentielle de départ est  $u(x) = e^{-x}$  ce qui nous permet de comparer à la figure 4.2, les solutions analytique et numérique pour  $N = 2$ . Ici encore, la comparaison est excellente sur  $u$  et un peu moins convaincante sur  $u'(x)$ . La situation s'améliore cependant très rapidement lorsque  $N$  augmente. ◀

### ► Exemple 4.3

On considère maintenant un problème bidimensionnel dont la géométrie est le carré de côtés de longueur 1 et les conditions aux limites sont décrites à la figure 4.3. On doit résoudre l'équation aux dérivées partielles :

$$\begin{cases} -\nabla^2 u = 10 & \text{dans } \Omega \\ u = g & \text{sur } \Gamma \text{ (} g \text{ décrite à la figure 4.3)} \end{cases}$$

Des conditions essentielles étant prescrites sur toute la frontière, on choisit  $V = H_0^1(\Omega)$  et comme toujours, on multiplie par  $w \in V$  et on intègre sur le domaine  $\Omega$ . On obtient :

$$\int_{\Omega} \nabla u \cdot \nabla w dv = \int_{\Omega} 10w dv$$

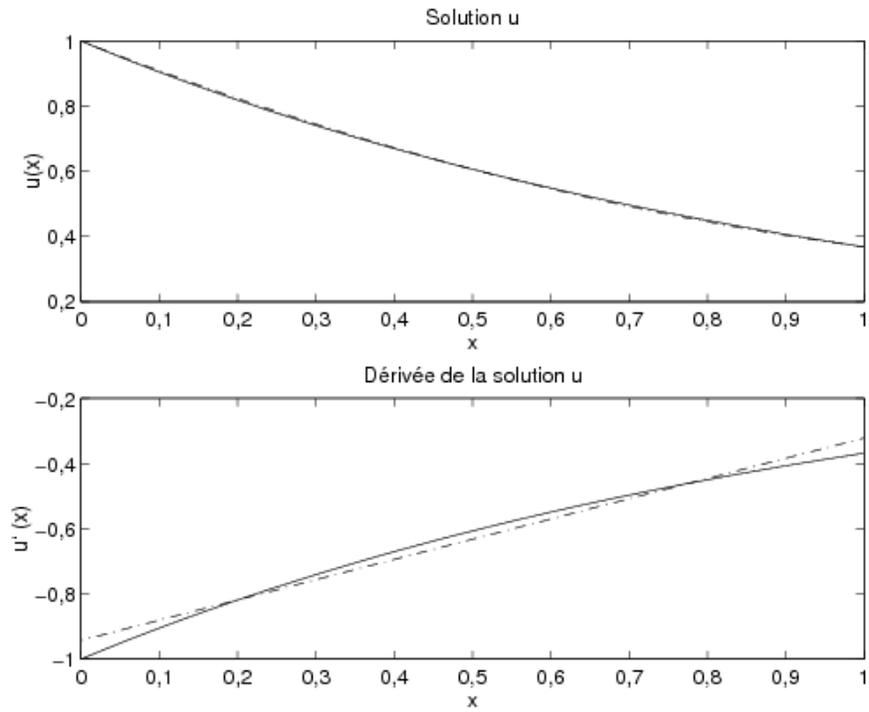


Figure 4.2 – Méthode de Ritz ( $N = 2$ )

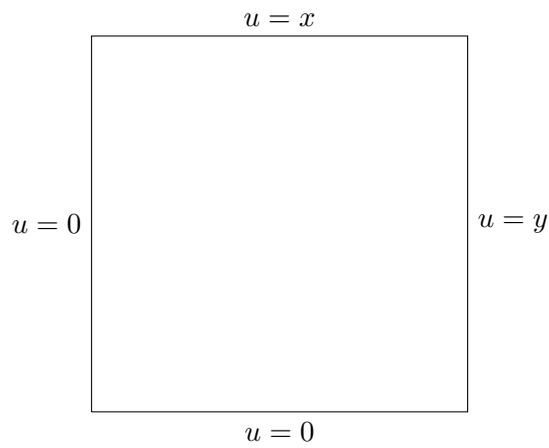


Figure 4.3 – Géométrie et conditions aux limites

On effectue le relèvement des conditions aux limites en posant  $u = \delta_u + u_g$ . Le choix de  $u_g$  est ici plus délicat malgré le fait que la géométrie du problème soit très simple. On peut prendre par exemple la fonction  $u_g(x, y) = xy$ . En remplaçant, on trouve la formulation variationnelle :

♠ trouver  $\delta_u \in V$  telle que :

$$\int_{\Omega} \nabla \delta_u \cdot \nabla w \, dv = \int_{\Omega} 10w \, dv - \int_{\Omega} \nabla u_g \cdot \nabla w \, dv$$

Il nous reste à construire les fonctions de Ritz  $\phi$ . Parmi les choix possibles, on peut prendre les  $N = m^2$  fonctions de la forme :

$$\phi_i(x, y) = \sin(i_1 \pi x) \sin(i_2 \pi y)$$

que l'on ordonne de la façon suivante :

$$\begin{aligned} \phi_1(x, y) &= \sin(\pi x) \sin(\pi y) \\ \phi_2(x, y) &= \sin(\pi x) \sin(2\pi y) \\ &\vdots \\ \phi_m(x, y) &= \sin(\pi x) \sin(m\pi y) \\ \phi_{m+1}(x, y) &= \sin(2\pi x) \sin(\pi y) \\ \phi_{m+2}(x, y) &= \sin(2\pi x) \sin(2\pi y) \\ &\vdots \\ \phi_{2m}(x, y) &= \sin(2\pi x) \sin(m\pi y) \\ &\vdots \\ \phi_{m^2}(x, y) &= \sin(m\pi x) \sin(m\pi y) \end{aligned}$$

On a ainsi une relation sur les indices de la forme :

$$i = (i_1 - 1)m + i_2 \quad 1 \leq i_1, i_2 \leq m$$

Ces fonctions s'annulent sur la frontière du domaine et vérifient une propriété d'orthogonalité très particulière :

$$a_{ij} = \int_{\Omega} \nabla \phi_j \cdot \nabla \phi_i \, dv = \begin{cases} 0 & \text{si } i \neq j \\ \frac{\pi^2}{4}(i_1^2 + i_2^2) & \text{si } i = j \end{cases}$$

La matrice  $A$  est donc diagonale ce qui est toutefois exceptionnel. De plus, on vérifie non sans peine que :

$$\begin{aligned} f_i &= \int_{\Omega} 10\phi_i \, dv - \int_{\Omega} \nabla u_g \cdot \nabla \phi_i \, dv = 10 \int_{\Omega} \phi_i \, dv - \int_{\Omega} (y, x) \cdot \nabla \phi_i \, dv \\ &= \frac{10}{\pi^2} \left( \frac{(-1)^{i_1} - 1}{i_1} \right) \left( \frac{(-1)^{i_2} - 1}{i_2} \right) \end{aligned}$$

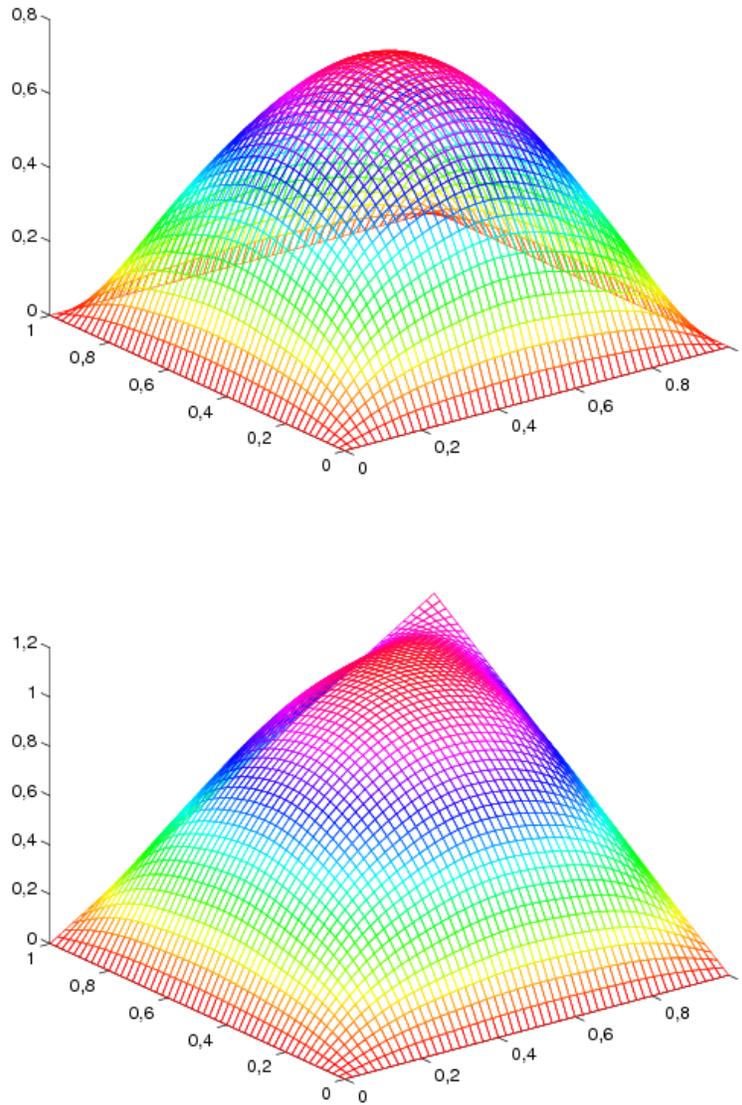


Figure 4.4 – Méthode de Ritz en dimension 2 : fonctions  $\delta_u(x)$  et  $u(x)$

Le système linéaire résultant est bien sûr très simple à résoudre. En prenant  $N = 5^2 = 25$ , on a obtenu le résultat de la figure 4.4 pour la fonction  $\delta_u$  qui rappelons-le, s'annule sur la frontière du domaine. La fonction  $u = \delta_u + u_g = \delta_u + xy$  est illustrée à la figure 4.4. ◀

Nous terminons ce chapitre par quelques remarques générales sur la méthode de Ritz. Les principales faiblesses de cette méthode sont :

- la construction de la fonction de relèvement  $u_g$  qui doit vérifier les conditions aux limites essentielles prescrites, ce qui peut être difficile particulièrement en dimension 2 ou 3 ;
- le choix et la construction des fonctions de Ritz  $\phi_i$  qui doivent être linéairement indépendantes et vérifier les conditions essentielles homogènes ;
- le calcul des coefficients  $a_{ij}$  et  $f_i$  du système linéaire qui peut nécessiter l'évaluation d'intégrales sur des domaines complexes. On peut éventuellement recourir aux méthodes d'intégration numériques si le besoin s'en fait sentir.
- Les inconnues  $u_i$  du système linéaire n'ont pas de signification physique particulière.

Les remarques précédentes sont particulièrement justifiées en dimension 2 ou 3 où les géométries peuvent être extrêmement variées et complexes. Il est alors pratiquement impossible de construire les différentes fonctions nécessaires. Cela est dû à l'approche globale de la méthode de Ritz, en ce sens que l'on cherche à construire les fonctions  $u_g$  et  $\phi_i$  en considérant tout le domaine  $\Omega$ . La méthode des éléments finis contourne cette difficulté par une approche plus localisée et une construction plus systématique des fonctions  $\phi_i$ .

### 4.3 Exercices

Déterminer pour les problèmes suivants une famille de fonctions  $\phi_i$  permettant d'appliquer la méthode de Ritz et obtenir la fonction de relèvement des conditions aux limites essentielles.

En dimension 1 :

1.

$$\begin{cases} -\nu_2 \frac{d^2 u}{dx^2} = f & \text{dans } ]0, 1[ \\ u(0) = u(1) = 0 \end{cases}$$

2.

$$\begin{cases} \nu_1 u - \nu_2 \frac{d^2 u}{dx^2} = f & \text{dans } ]0, 1[ \\ \frac{du}{dx}(0) = \frac{du}{dx}(1) = 0 \end{cases}$$

3.

$$\begin{cases} -\nu_2 \frac{d^2 u}{dx^2} = f & \text{dans } ]0, 1[ \\ u(0) = a, \quad u(1) = b \end{cases}$$

4.

$$\begin{cases} -\frac{d}{dx} \left( q(x) \frac{du}{dx} \right) = f & \text{dans } ]0, 1[ \\ u(0) = a, \quad q(1) \frac{du}{dx}(1) = p \end{cases}$$

5.

$$\begin{cases} \frac{d^2}{dx^2} \left( q(x) \frac{d^2 u}{dx^2} \right) = f & \text{dans } ]0, L[ \\ u(0) = a, \quad u(L) = b \\ q(0) \frac{d^2 u}{dx^2}(0) = c, \quad q(L) \frac{d^2 u}{dx^2}(L) = d \end{cases}$$

6.

$$\begin{cases} \frac{d^2}{dx^2} \left( q(x) \frac{d^2 u}{dx^2} \right) = f & \text{dans } ]0, L[ \\ u(0) = a, \quad u(L) = b \\ u'(0) = 0, \quad q(L) \frac{d^2 u}{dx^2}(L) = d \end{cases}$$

En dimension 2 :

$$7. \quad \left\{ \begin{array}{ll} -\nabla \cdot (q(\mathbf{x}) \nabla u) = f & \text{dans le carré } ]0, 1[^2 \\ u = 1 & \text{sur les côtés } x = 0 \text{ et } y = 0 \\ q(\mathbf{x}) \frac{\partial u}{\partial \mathbf{n}} = 2 & \text{sur les autres côtés} \end{array} \right.$$

8. On souhaite résoudre le problème suivant par la méthode de Ritz :

$$\frac{d^2}{dx^2} \left( EI \frac{d^2 u(x)}{dx^2} \right) = f(x) \text{ dans } ]0, 1[,$$

où  $EI$  est une constante. Proposer des choix de fonctions  $u_g(x)$  et  $\phi_i(x)$  permettant de résoudre ce problème avec les conditions aux limites suivantes :

$$u(0) = 1, \quad u'(0) = 2, \quad u(1) = 4, \quad EI \frac{d^2 u(1)}{dx^2} = M$$

(Utiliser de préférence des fonctions polynômiales).

9. Résoudre par la méthode de Ritz le problème suivant :

$$\left\{ \begin{array}{ll} -\nabla \cdot k \nabla T = q & \text{dans le carré } ]0, 1[ \times ]0, 1[ \\ T = 0 & \text{sur les côtés } x = 1 \text{ et } y = 1 \\ \frac{\partial T}{\partial \mathbf{n}} = 0 & \text{sur les autres côtés} \end{array} \right.$$

où  $k$  et  $q$  sont des constantes. On prendra une seule fonction de Ritz :

$$\phi_1(x) = (1 - x^2)(1 - y^2)$$

10. Résoudre par la méthode de Ritz le problème suivant :

$$\left\{ \begin{array}{ll} -\frac{d}{dx} \left( (x^2 + 1) \frac{du}{dx} \right) = 2 - 2x + 6x^2 & \text{dans l'intervalle } ]0, 1[ \\ u(0) = 1 \\ u'(1) = -1 \end{array} \right.$$

- a) Obtenir la formulation variationnelle ;
- b) Choisir les fonctions de Ritz appropriées ;

- c) Construire le relèvement des conditions aux limites essentielles ;
  - d) Obtenir les coefficients  $a_{ij}$  et  $f_i$  du système linéaire correspondant ;
  - e) Résoudre le système en utilisant deux fonctions de Ritz.
11. On suppose que l'on a un problème variationnel de la forme :

$$a(u, w) = l(w)$$

où la forme bilinéaire **symétrique**  $a$  et la forme linéaire  $l$  vérifient les hypothèses du théorème de Lax-Milgram dans un espace de Hilbert  $V$ .

Montrer que si on utilise la méthode de Ritz pour discrétiser ce problème, on obtient une matrice  $A$  définie positive.

**Rappel :** Une matrice symétrique  $A$  est dite définie positive si quel que soit le vecteur colonne (non nul)  $\mathbf{x} = [x_1, x_2, x_3, \dots, x_n]^T$ , on a :

$$\mathbf{x}^T A \mathbf{x} = (A \mathbf{x}) \cdot \mathbf{x} = \sum_{i,j=1}^n A_{ij} x_i x_j > 0$$



## Chapitre 5

# Éléments finis unidimensionnels

Nous allons maintenant établir comment construire les fonctions  $\phi_i(x)$  de la méthode de Ritz de manière efficace, et ce sur des domaines de forme quelconque. Cette efficacité proviendra de l'introduction de formes géométriques simples (des sous-intervalles en dimension 1) nommés *éléments* qui permettent une construction locale de ces fonctions. C'est certainement en dimension 2 ou 3 que l'introduction de ces éléments prend toute son importance car les géométries sont de toute évidence beaucoup plus complexes. Nous commencerons tout de même le développement en dimension 1 en tâchant d'utiliser une présentation la plus générale possible de sorte que le passage en dimension supérieure soit direct.

Nous choisissons une approche qui se veut la plus pédagogique possible. Nous présenterons aussi quelques aspects informatiques en indiquant quelques tableaux nécessaires à la bonne mise en oeuvre d'une méthode d'éléments finis. Nous ne chercherons cependant pas à obtenir une présentation optimale sur le plan informatique. Ce n'est pas le but de cet ouvrage. Le lecteur doit rester conscient qu'il existe plusieurs façons de présenter la méthode des éléments finis et que celle que nous avons retenue a l'avantage d'être relativement simple et suffisamment générale.

### 5.1 Équations différentielles d'ordre 2

#### 5.1.1 Problème type

Pour fixer les idées, nous commencerons par la résolution d'un problème classique d'une équation différentielle d'ordre 2 de la forme :

$$\begin{cases} p(x)u - \frac{d}{dx} \left( q(x) \frac{du}{dx} \right) = r(x) & \text{dans } ]0, L[ \\ u(0) = c, \quad q(L) \frac{du}{dx}(L) = d \end{cases} \quad (5.1)$$

On suppose connues les fonctions  $p(x)$  et  $q(x)$  de même que les constantes  $c$  et  $d$ .

Si dans cette équation on prend  $p(x) = 0$ , la variable dépendante  $u(x)$  peut entre autres choses désigner la déformation longitudinale d'une tige métallique de longueur  $L$ . À ce moment,  $q(x) = EA$ , où  $E$  est le module de Young et  $A$  l'aire de la section de la tige qui peut être variable. Enfin,  $r(x)$  est une force de contact sur la surface de la tige et  $d$  est une force axiale appliquée à l'une de ses extrémités. On peut aussi interpréter  $u(x)$  comme une température, une pression hydrostatique, un potentiel, etc., suivant le domaine qui nous intéresse (voir par exemple Reddy, réf. [31]). Le problème 5.1 est donc suffisamment général pour couvrir bon nombre d'applications.

Sur le plan théorique, si on suppose que la fonction  $p(x)$  est positive ou nulle et que  $0 < q_1 \leq q(x) \leq q_2$  dans l'intervalle  $]0, L[$ , il est facile de s'assurer que les hypothèses du théorème de Lax-Milgram sont vérifiées. La solution, à un relèvement des conditions essentielles près, est dans l'espace :

$$V = \{w \in H^1(]0, L[) \mid w(0) = 0\}$$

Rappelons que pour cette équation différentielle d'ordre 2, la seule condition essentielle est l'imposition de la variable  $u(x)$  à la frontière. Dans cet exemple,  $u$  n'est imposée qu'en  $x = 0$  d'où cette définition de l'espace  $V$ .

Il est donc nécessaire de construire des fonctions  $\phi_i(x)$  dans la méthode de Ritz qui soient dans  $V$ . Comme nous l'avons vu au chapitre 2 et puisque nous utiliserons des approximations polynômiales, il suffit de s'assurer que ces fonctions soient continues (et s'annulent en  $x = 0$  dans ce cas précis).

On pose ensuite  $u = \delta_u + u_g$  et la formulation variationnelle (laissée en exercice) est :

♠ trouver  $\delta_u \in V$  telle que :

$$a(\delta_u, w) = l(w) - a(u_g, w) \quad \forall w \in V \quad (5.2)$$

où :

$$a(\delta_u, w) = \int_0^L (p(x)\delta_u(x)w(x) + q(x)\delta_u'(x)w'(x)) dx$$

et :

$$l(w) = d w(L) + \int_0^L r(x)w(x)dx$$

La méthode de Ritz consiste à poser :

$$u(x) = u_g(x) + \delta_u(x) = u_g(x) + \sum_{j=1}^n \delta_{u_j} \phi_j(x)$$

et nécessite la construction d'un système linéaire global de la forme :

$$A\delta_U = F$$

où :

$$a_{ij} = a(\phi_j(x), \phi_i(x)) \quad \text{et} \quad f_i = l(\phi_i(x)) - a(u_g(x), \phi_i(x))$$

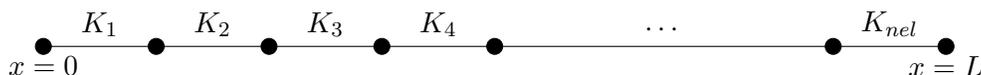


Figure 5.1 – Maillage en dimension 1

La fonction  $u_g(x)$  est le relèvement des conditions aux limites essentielles et doit donc vérifier dans ce cas  $u_g(0) = c$  et doit de plus appartenir à l'espace  $H^1(]0, L[)$ .

#### ◆ Remarque 5.1

Il est souvent utile, en particulier pour les problèmes non linéaires que nous rencontrerons au chapitre 8, de considérer la fonction  $\delta_u(x)$  comme une correction à une solution existante  $u_g(x)$  obtenue préalablement et vérifiant bien sûr les conditions essentielles imposées. La fonction  $u_g(x)$  pourra provenir d'un calcul précédent ou tout simplement d'une itération précédente. ◆

Voyons maintenant comment construire les fonctions de Ritz  $\phi_i(x)$ . Cette construction se fera en plusieurs étapes qui constituent ce que l'on appelle la *méthode de Ritz-Galerkin*.

### 5.1.2 Le maillage

#### Les éléments

Considérons une partition de l'intervalle  $]0, L[$  comme celle de la figure 5.1. Cette partition en sous-intervalles constitue le *maillage*. Les sous-intervalles  $K_i$  sont appelés *éléments* dont le nombre total est noté  $nel$ . La frontière des éléments correspond aux cercles en trait gras. Les éléments peuvent être de longueur variable.

#### Les noeuds

Sur chaque élément, on identifie  $n_g^K$  *noeuds géométriques* qui permettent de définir la géométrie de l'élément en question. En dimension 1, les noeuds géométriques sont les bornes de l'élément et alors  $n_g^K = 2$ . On notera  $K = [x_1^K, x_2^K]$  un élément générique et  $h^K$  sa taille qui en dimension 1, est tout simplement la longueur  $h^K = x_2^K - x_1^K$  (on s'assurera bien sûr que  $x_2^K > x_1^K$  de sorte que  $h^K$  soit strictement positif). Ce sont ces noeuds géométriques qui définissent localement la géométrie de l'élément et globalement, celle du domaine en entier. En dimension 1, cela paraît futile mais en dimension supérieure, ce sera fondamental.

Dans chaque élément  $K$  on identifie également (voir la figure 5.2) un nombre  $n_c^K$  de *noeuds de calcul* souvent appelés *noeuds d'interpolation* où les variables essentielles du problème seront éventuellement calculées. Ces noeuds d'interpolation peuvent coïncider ou non avec les noeuds géométriques. L'ensemble des noeuds de l'élément comprend donc les noeuds géométriques et les noeuds de calcul. On note  $n_t^K$  le nombre total de noeuds de l'élément  $K$  et  $(x_1^K, x_2^K, \dots, x_{n_t^K}^K)$  les

Élément  $K$ Figure 5.2 – Noeuds géométriques et de calcul de l'élément  $K$ 

noeuds de l'élément  $K$  en commençant par les noeuds géométriques, suivis des autres noeuds de calcul.

## ◆ Remarque 5.2

Pour éviter d'alourdir inutilement l'exposé, nous supposons que les noeuds de calcul comprennent les noeuds géométriques, même si cette hypothèse n'est absolument pas nécessaire. Le nombre total de noeuds de l'élément est donc  $n_t^K = n_c^K$ . On rencontre le cas où les noeuds de calcul diffèrent totalement des noeuds géométriques notamment pour les éléments dits non conformes (voir Ciarlet, réf. [9]). ◆

On place les coordonnées de tous les noeuds (géométriques et/ou de calcul) dans un tableau que nous notons *coor* de longueur égale au nombre de noeuds total *nnoeuds* du maillage. On définit un *tableau de connectivité des noeuds* (notée *connec*) comprenant *nel* lignes. À chacune de ces lignes, on retrouve les numéros des  $n_c^K$  noeuds de l'élément.

## Les degrés de liberté

On associe à chaque noeud de calcul de l'élément une ou plusieurs inconnues appelées *degrés de liberté* (ddl) suivant que le problème possède une ou plusieurs variables essentielles. Il est même possible qu'à un noeud donné, aucun degré de liberté ne soit attribué ou qu'en 2 noeuds d'un même élément, un nombre différent de degrés de liberté ne soit associés. On s'assure ainsi d'une grande flexibilité au niveau de l'implantation de la méthode. On note  $n_d^K$  le nombre de degrés de liberté de l'élément. Très souvent,  $n_d^K$  est un multiple de  $n_c^K$ . Certains degrés de liberté seront communs à 2 (ou plus) éléments. Le nombre total de degrés de liberté du domaine sera noté *nddl*. Les degrés de liberté sont donc les inconnues de notre problème. Paradoxalement, certains degrés de liberté sont fixes puisqu'imposés par les conditions essentielles du problème. On notera :

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{nddl} \end{bmatrix} = \begin{bmatrix} U^I \\ U^C \end{bmatrix} \quad \text{et} \quad \delta_U = \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \vdots \\ \delta_{u_{nddl}} \end{bmatrix} = \begin{bmatrix} \delta_U^I \\ 0 \end{bmatrix} \quad (5.3)$$

les vecteurs globaux des degrés de liberté pour  $u(x)$  et  $\delta_u(x)$  qui sont de dimension  $nddl$ . Le vecteur  $U$  se décompose en 2 parties distinctes :  $U^C$  est la partie de  $U$  qui est connue (d'où l'indice  $C$ ) car imposée par les conditions aux limites essentielles du problème. La partie correspondante dans le vecteur de correction  $\delta_U$  est par conséquent nulle. Le reste est noté  $U^I$  (et  $\delta_U^I$  pour la correction) ( $I$  pour inconnu) et sera éventuellement calculé. Nous reviendrons sur cette partition de  $U$  lors de l'imposition des conditions aux limites (voir la section 5.1.8) et lors de la résolution du système global (voir la section 5.1.9).

À chaque degré de liberté du domaine, on attribue une fonction de Ritz  $\phi(x)$  et, au noeud de calcul correspondant, on calculera une approximation de la variable essentielle associée. Chaque degré de liberté est relié à un noeud de calcul mais inversement, un noeud de calcul peut être associé à plusieurs degrés de liberté. On construira ainsi un système linéaire (ou non linéaire suivant que le problème de départ est linéaire ou non) de dimension  $nddl$  sur  $nddl$ .

### Numérotation des degrés de liberté

Le tableau de connectivité des noeuds *connec* est très utile mais ne suffit pas complètement puisque nous avons mentionné qu'à un noeud, on peut associer 1 ou plusieurs degrés de liberté. Il faut aussi introduire ce que nous appellerons un *tableau de numérotation* des degrés de liberté. Pour ce faire, on assigne en tout premier lieu un numéro à chaque degré de liberté du problème. Cette numérotation fera en sorte que les degrés de liberté où on impose une condition essentielle seront numérotés en dernier.

Pour y arriver, on construit un tableau de numérotation noté *numer* qui contient pour chaque noeud les numéros des degrés de liberté qui lui sont associés. C'est donc un tableau de dimension  $nnoeuds$  multipliée par le nombre de degrés de liberté par noeud qui peut être variable. Les étapes de construction du tableau *numer* sont les suivantes :

1. pour chacune des variables essentielles du problème, on identifie les numéros des noeuds où cette variable est imposée ;
2. on parcourt les noeuds en numérotant tous les degrés de liberté associés à chacun des ces noeuds et qui ne sont pas imposés comme condition aux limites essentielle du problème ;
3. on parcourt de nouveau les noeuds et on numérote à la suite tous les autres degrés de liberté c.-à-d. tous les degrés de liberté qui sont imposés comme condition aux limites essentielle du problème.

De cette manière, on décompose la matrice du système global et le vecteur des degrés de liberté du problème en 2 parties consécutives où on regroupe au début les véritables inconnues du système (le vecteur  $U^I$  que l'on devra calculer) et à la fin les valeurs des variables essentielles imposées (le vecteur  $U^C$ ) tel qu'indiqué à la relation 5.3.

Enfin, une fois la numérotation obtenue, on construit le *tableau d'adressage* noté *adres* qui pour chaque élément fournit les numéros des degrés de liberté associés à chaque noeud de calcul de l'élément. Chaque ligne de ce tableau correspond à un élément et sera appelée le *vecteur d'adressage* de cet élément. Le tableau d'adressage peut être construit une fois pour toutes grâce aux tableaux

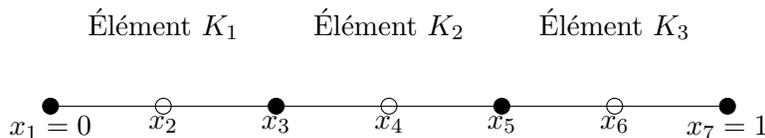


Figure 5.3 – Maillage :  $nel = 3, n_g^K = 2, n_c^K = 3$

*connec* et *numer*. Il est cependant souvent préférable de construire le vecteur d’adressage de chaque élément au besoin. L’exemple qui suit illustre les différentes étapes et tableaux que nous venons d’introduire.

► **Exemple 5.1**

La situation est illustrée à la figure 5.3 pour notre problème type. Ce maillage très simple du domaine  $]0, 1[$  (on choisit ici  $L = 1$ ) comporte  $nel = 3$  éléments dont les frontières sont identifiées par des cercles en trait gras. Sur chaque élément on a identifié 3 noeuds de calcul ( $n_c^K = 3$ ) comprenant les extrémités (les noeuds géométriques) et le point-milieu de l’élément. Certains noeuds sont communs à 2 éléments comme par exemple  $x_3$  (commun aux éléments  $K_1$  et  $K_2$ ) et  $x_5$  (commun à  $K_2$  et  $K_3$ ).

Dans cet exemple, on a  $nnoeuds = 7$  et le tableau *coor* est :

<b>Coordonnées des noeuds</b>		
Tableau <i>coor</i>		
$x_1$	$x_2$	$x_3$
0,0000	0,0000	0,0000
0,1667	0,0000	0,0000
0,3333	0,0000	0,0000
0,5000	0,0000	0,0000
0,6667	0,0000	0,0000
0,8333	0,0000	0,0000
1,0000	0,0000	0,0000

Bien entendu, les coordonnées  $x_2$  et  $x_3$  sont inutiles pour un problème unidimensionnel. Le tableau de connectivité des noeuds *connec* prend la forme :

<b>Numéros des noeuds géométriques et de calcul</b>			
Tableau <i>connec</i>			
	Noeuds géométriques		Autres noeuds
Élément	Noeud #1	Noeud #2	Noeud #3
1	1	3	2
2	3	5	4
3	5	7	6

Chaque ligne de ce tableau correspond à un élément différent. Rappelons encore ici la convention établie de numérotter en premier lieu les noeuds géométriques, et ce pour chaque élément. Dans cet exemple, il y a plus de noeuds de calcul que de noeuds géométriques mais ce n'est pas toujours le cas. Les deux tableaux précédents ainsi que le tableau *coord* contenant les coordonnées de tous les noeuds nous permettent de définir complètement la géométrie du domaine et contiennent aussi des informations que nous utiliserons un peu plus loin.

Enfin, il faut souligner que le tableau *connec* ne contient que les numéros des noeuds (de calcul et/ou géométriques). Pour obtenir les coordonnées de l'élément  $K_2$  par exemple, il faut utiliser le tableau *coord* et poser :

$$\begin{aligned} K_2 &= [x_1^{K_2}, x_2^{K_2}] \\ &= [\text{coord}(\text{connec}(2, 1)), \text{coord}(\text{connec}(2, 2))] \\ &= [\text{coord}(3), \text{coord}(5)] \\ &= [0,3333, 0,6667] \end{aligned}$$

et le deuxième élément de ce maillage est donc l'intervalle  $[1/3, 2/3]$ . De plus, les coordonnées des noeuds de calcul de ce même élément sont :

$$\begin{aligned} (\text{coord}(\text{connec}(2, i)), i = 1, 2, \dots, n_c^K) &= \\ (\text{coord}(3), \text{coord}(5), \text{coord}(4)) &= \\ (0,3333, 0,6667, 0,5000) & \end{aligned}$$

Dans cet exemple (voir les équations 5.1), la variable essentielle  $u$  n'est imposée qu'en  $x = 0$  au noeud 1. Le degré de liberté associé au noeud 1 sera donc numéroté en dernier. Dans un premier temps, on laisse tomber les degrés de liberté imposés. Une première numérotation possible serait alors :

Numérotation	
Tableau <i>numer</i>	
Noeud	$u$
1	?
2	1
3	2
4	3
5	4
6	5
7	6

On reparcourt ensuite les noeuds pour numérotter les degrés de liberté fixés par les conditions aux limites essentielles. On obtient :

Numérotation	
Tableau <i>numer</i>	
Noeud	<i>u</i>
1	7
2	1
3	2
4	3
5	4
6	5
7	6

Chaque ligne de ce tableau correspond à un noeud de calcul du domaine. Dans le cas où on associe plusieurs degrés de liberté à un même noeud, on ajoute des colonnes supplémentaires à ce tableau. On peut même à la limite avoir un tableau dont les lignes ne sont pas toutes de même longueur dans le cas où le nombre de degrés de liberté associés n'est pas le même d'un noeud à l'autre.

On peut maintenant construire le tableau d'adressage dont chaque ligne correspond au vecteur d'adressage d'un élément  $K$  et fournit les numéros des degrés de liberté qui lui sont associés. On l'obtient en posant :

$$adres(K, i) = numer(connec(K, i)), i = 1, 2, \dots, n_c^K$$

Dans notre exemple, on aurait :

Numéros des ddl			
Tableau <i>adres</i>			
Élément	Ddl #1	Ddl #2	Ddl #3
1	7	2	1
2	2	4	3
3	4	6	5

Chaque ligne correspond à un élément. On peut facilement reconstruire ce tableau à l'aide des tableaux *connec* et *numer*. En effet, pour la première ligne (le premier élément), on a :

$$\begin{aligned} adres(1, 1) &= numer(connec(1, 1)) = numer(1) = 7 \\ adres(1, 2) &= numer(connec(1, 2)) = numer(3) = 2 \\ adres(1, 3) &= numer(connec(1, 3)) = numer(2) = 1 \end{aligned}$$

Il en est de même pour chaque élément. Rappelons qu'il n'est pas nécessaire de garder en mémoire ce tableau puisqu'on peut le reconstruire très facilement à partir des tableau *connec* et *numer* lorsque l'on en a besoin. ◀

### 5.1.3 Formulation variationnelle élémentaire

Cette étape consiste à obtenir une formulation variationnelle sur un élément quelconque  $K$ . On utilise donc la méthode de Ritz comme au chapitre précédent, à l'exception notable près que l'on intègre sur un élément  $K = [x_1^K, x_2^K]$  plutôt que sur le domaine  $\Omega$  au complet. On obtient alors, après une intégration par parties :

$$\int_K \left( p(x)u(x)w(x) + q(x)\frac{du}{dx}\frac{dw}{dx} \right) dx = \int_K r(x)w(x)dx + q(x)\frac{du}{dx}w(x) \Big|_{x_1^K}^{x_2^K}$$

ou encore :

$$\int_{x_1^K}^{x_2^K} \left( p(x)u(x)w(x) + q(x)\frac{du}{dx}\frac{dw}{dx} \right) dx = \int_{x_1^K}^{x_2^K} r(x)w(x)dx + s_{12}^K w(x_2^K) + s_{11}^K w(x_1^K)$$

où on a introduit les *variables secondaires* :

$$s_{11}^K = -q(x_1^K)\frac{du}{dx}(x_1^K) \quad \text{et} \quad s_{12}^K = q(x_2^K)\frac{du}{dx}(x_2^K)$$

Ces variables secondaires correspondent aux conditions aux limites naturelles aux extrémités de l'élément. La notation à 2 indices indique que  $s_{11}^K$  est la valeur de la première variable secondaire à la première extrémité de l'élément tandis que  $s_{12}^K$  est la valeur de la première variable secondaire à la deuxième extrémité de l'élément  $K$ . On prévoit ainsi le cas où il y aura plusieurs variables secondaires définies aux bornes des éléments ce qui sera le cas pour les problèmes d'ordre 4. On qualifie ces variables de secondaires par opposition aux variables essentielles dites *variables primaires* suivant la notation de Reddy, réf. [31].

Puisque nous avons convenu de travailler en correction  $\delta_u$ , on pose encore  $u(x) = \delta_u(x) + u_g(x)$  et on a ainsi la formulation variationnelle élémentaire :

$$\begin{aligned} \int_{x_1^K}^{x_2^K} \left( p(x)\delta_u(x)w(x) + q(x)\frac{d\delta_u}{dx}\frac{dw}{dx} \right) dx &= \int_{x_1^K}^{x_2^K} r(x)w(x)dx + s_{12}^K w(x_2^K) + s_{11}^K w(x_1^K) \\ &\quad - \int_{x_1^K}^{x_2^K} \left( p(x)u_g(x)w(x) + q(x)\frac{du_g}{dx}\frac{dw}{dx} \right) dx \end{aligned}$$

et l'on retrouve bien ainsi la forme 5.2.

Nous ne sommes pas en mesure d'imposer les conditions aux limites puisque nous sommes sur un élément qui n'est pas forcément situé à la frontière du domaine. L'idée consiste maintenant à appliquer la méthode de Ritz sur l'élément  $K$ . On utilise donc sur chaque élément  $K$  une relation de la forme :

$$u(x) |_K \simeq \delta_u^K(x) + u_g^K(x)$$

où l'indice supérieur  $K$  désigne la restriction à l'élément  $K$ . On pose ensuite :

$$\delta_u^K(x) = \sum_{j=1}^{n_d^K} \delta_{u_j}^K \psi_j^K(x) \quad \text{et} \quad u_g^K(x) = \sum_{j=1}^{n_d^K} u_{g_j}^K \psi_j^K(x) \quad (5.4)$$

où  $n_d^K$  est le nombre total de degrés de liberté associés à la variable essentielle  $\delta_u(x)$  sur l'élément  $K$ . Les  $\delta_{u_j}^K$  sont les valeurs des degrés de liberté de l'élément et sont inconnus (sauf là où des conditions de Dirichlet seront imposées). Les  $u_{g_j}^K$  sont les valeurs nodales du relèvement des conditions aux limites et sont supposées connues. Nous verrons comment les obtenir un peu plus loin.

Comme cela est le cas ici, on aura très souvent  $n_d^K = n_c^K$  mais ce n'est aucunement nécessaire. Cette situation provient du fait qu'il n'y a qu'une seule variable essentielle (primaire) et que l'on associe un seul degré de liberté à chaque noeud de calcul. Si on remplace dans la formulation variationnelle, on obtient :

$$\begin{aligned} \sum_{j=1}^{n_d^K} \delta_{u_j}^K \int_{x_1^K}^{x_2^K} \left( p(x) \psi_j^K(x) w(x) + q(x) \frac{d\psi_j^K}{dx} \frac{dw}{dx} \right) dx &= \int_{x_1^K}^{x_2^K} r(x) w(x) dx + s_{12}^K w(x_2^K) + s_{11}^K w(x_1^K) \\ &\quad - \int_{x_1^K}^{x_2^K} \left( p(x) u_g^K(x) w(x) + q(x) \frac{du_g^K}{dx} \frac{dw}{dx} \right) dx \end{aligned}$$

Les fonctions  $\psi_j^K(x)$  sont appelées *fonctions d'interpolation de l'élément  $K$*  et ne sont définies que sur  $K$  et non sur le domaine  $\Omega$  au complet. Pour obtenir un système linéaire, il suffit de prendre successivement  $w(x) = \psi_i^K(x)$ ,  $i = 1, 2, \dots, n_d^K$ . On obtient alors le *système élémentaire*  $n_d^K$  sur  $n_d^K$  suivant :

$$A^K \delta_U^K = F^K + S^K \quad (5.5)$$

où :

$$\begin{aligned} a_{ij}^K &= \int_{x_1^K}^{x_2^K} \left( p(x) \psi_j^K(x) \psi_i^K(x) + q(x) \frac{d\psi_j^K}{dx} \frac{d\psi_i^K}{dx} \right) dx \\ f_i^K &= \int_{x_1^K}^{x_2^K} r(x) \psi_i^K(x) dx - \int_{x_1^K}^{x_2^K} \left( p(x) u_g^K(x) \psi_i^K(x) + q(x) \frac{du_g^K(x)}{dx} \frac{d\psi_i^K(x)}{dx} \right) dx \\ s_i^K &= s_{11}^K \psi_i^K(x_1^K) + s_{12}^K \psi_i^K(x_2^K) \end{aligned}$$

La *matrice élémentaire*  $A^K$  (de coefficients  $a_{ij}^K$ ) est souvent appelée *matrice de rigidité* faisant ainsi référence aux premières applications de la méthode des éléments finis dans le domaine des structures. Le vecteur  $\delta_U^K$  (de coefficients  $\delta_{u_i}^K$ ) est appelé *vecteur des degrés de liberté élémentaires*. Bien que cela ne soit pas absolument nécessaire, nous avons séparé le terme de droite du système élémentaire en 2 parties  $F^K$  (de coefficients  $f_i^K$ ) et  $S^K$  (de coefficients  $s_i^K$ ). Cela permet d'isoler

la contribution  $S^K$  des variables secondaires qui nécessiteront un traitement particulier lors de l'imposition des conditions aux limites. Le terme de droite au complet  $(F^K + S^K)$  est le *vecteur des sollicitations élémentaires*.

Une telle formulation requiert la construction des fonctions  $\psi_i^K$  sur chaque élément  $K$  ce qui constitue une procédure assez lourde. De plus, on a souvent recours à l'intégration numérique pour évaluer les coefficients du système élémentaire. Comme nous le verrons un peu plus loin, cela nécessite la mise en mémoire de *points d'intégration* différents d'un élément à l'autre, ce qui exigerait beaucoup d'espace mémoire. Pour contourner cette difficulté, on introduit un élément  $\hat{K}$  dit de *référence* sur lequel on effectue toutes les intégrales nécessaires à l'évaluation du système élémentaire, et ce au moyen d'un changement de variables.

#### 5.1.4 Passage à l'élément de référence

En dimension 1, nous prendrons habituellement l'intervalle  $\hat{K} = [-1, 1]$  comme élément de référence. Le changement de variables de l'élément de référence  $\hat{K}$  à l'*élément réel*  $K$  est aussi appelé *transformation géométrique* et s'exprime sous la forme :

$$\begin{aligned}
 T^K : \quad \hat{K} &\rightarrow K \\
 [-1, 1] &\rightarrow [x_1^K, x_2^K] \\
 \xi &\rightarrow x = \frac{(x_1^K + x_2^K) + h^K \xi}{2} \\
 dx &= \frac{h^K}{2} d\xi
 \end{aligned} \tag{5.6}$$

Cette transformation géométrique est inversible et il en sera toujours ainsi, même en dimension supérieure pour les transformations linéaires. Toutefois, on peut aisément concevoir des transformations non linéaires de l'élément de référence et il faut alors être prudent car il est possible que dans certaines situations, la transformation ne soit pas inversible. On peut se référer à Dhatt-Touzot, réf. [13], à ce sujet.

C'est à cette étape que les tableaux *connec* et *coor* sont utiles car ils permettent de calculer les termes de la transformation  $T^K$  pour chaque élément. Ainsi, à chaque point  $\xi$  de l'élément de référence correspond un point  $x$  de l'élément  $K$  et inversement. Dans le cas présent la transformation

$T^K$  est facile à inverser et on a :

$$\begin{aligned} (T^K)^{-1} : \quad K &\rightarrow \hat{K} \\ [x_1^K, x_2^K] &\rightarrow [-1, 1] \\ x &\rightarrow \xi = \frac{2x - (x_1^K + x_2^K)}{h^K} \\ d\xi &= \frac{2}{h^K} dx \end{aligned} \tag{5.7}$$

On remarque que les extrémités de l'élément courant  $K$  sont envoyées sur les extrémités de l'élément de référence. On peut également obtenir cette transformation en introduisant les fonctions d'interpolation de Lagrange (voir l'annexe C ou Fortin, réf. [17]) sur l'élément de référence :

$$L_1(\xi) = \frac{1 - \xi}{2} \quad \text{et} \quad L_2(\xi) = \frac{1 + \xi}{2}$$

La transformation  $T^K$  s'écrit alors également :

$$x = x_1^K L_1(\xi) + x_2^K L_2(\xi) = \frac{(x_1^K + x_2^K) + h^K \xi}{2} \tag{5.8}$$

Cette procédure est plus générale et permet de définir différentes transformations de l'élément de référence vers l'élément  $K$ . Nous procéderons de cette façon en dimension supérieure à 1. On peut encore ici facilement construire de manière similaire des transformations non linéaires en prenant des fonctions de Lagrange de degré supérieur.

C'est *a priori* uniquement sur l'élément de référence que nous construirons des fonctions d'interpolation  $\hat{\psi}_i(\xi)$ . On définit ensuite les fonctions d'interpolation  $\psi_i^K(x)$  sur l'élément  $K$  par composition :

$$\psi_i^K(x) = \psi_i^K(T^K(\xi)) = \hat{\psi}_i(\xi) \quad \text{ou encore} \quad \hat{\psi}_i(\xi) = \hat{\psi}_i((T^K)^{-1}(x)) = \psi_i^K(x)$$

Les fonctions  $\psi_i^K(x)$  ne sont que rarement explicitées puisque, comme nous le verrons, nous n'en avons aucunement besoin.

Ainsi, la fonction d'interpolation  $\psi_i^K(x)$  prendra la même valeur que la fonction  $\hat{\psi}_i(\xi)$  au point  $x$  tel que  $x = T^K(\xi)$ . Les fonctions  $\hat{\psi}_i(\xi)$  sont par le fait même généralement indépendantes des éléments  $K$  et sont appelées *fonctions d'interpolation de l'élément de référence*. On a ainsi un seul ensemble de fonctions à construire et on transforme les dérivées par la formule de dérivation en chaîne :

$$\begin{aligned} \frac{d\psi_j^K}{dx} &= \frac{d\hat{\psi}_j}{d\xi} \frac{d\xi}{dx} = \frac{d\hat{\psi}_j}{d\xi} \frac{2}{h^K} \\ \frac{d^2\psi_j^K}{dx^2} &= \frac{d}{dx} \left( \frac{d\psi_j^K}{dx} \right) = \frac{d}{dx} \left( \frac{2}{h^K} \frac{d\hat{\psi}_j}{d\xi} \right) = \frac{d}{d\xi} \left( \frac{2}{h^K} \frac{d\hat{\psi}_j}{d\xi} \right) \frac{d\xi}{dx} = \frac{4}{(h^K)^2} \frac{d^2\hat{\psi}_j}{d\xi^2} \end{aligned} \tag{5.9}$$

Il convient de préciser comment on évalue  $u_g(x)$  et  $\frac{du_g(x)}{dx}$  dans le terme de droite. On utilise bien entendu l'expression 5.4 de sorte que :

$$u_g^K(x) = \sum_{j=1}^{n_d^K} u_{g_j}^K \psi_j^K(x) = \sum_{j=1}^{n_d^K} u_{g_j}^K \hat{\psi}_j(\xi) \quad \text{et} \quad \frac{du_g^K(x)}{dx} = \sum_{j=1}^{n_d^K} u_{g_j}^K \frac{d\psi_j^K(x)}{dx} = \frac{2}{h^K} \sum_{j=1}^{n_d^K} u_{g_j}^K \frac{d\hat{\psi}_j(\xi)}{d\xi}$$

Effectuons maintenant le changement de variables dans le système élémentaire dont les coefficients deviennent :

$$\begin{aligned} a_{ij}^K &= \int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) \frac{h^K}{2} d\xi \\ &\quad + \int_{-1}^1 q \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \left( \frac{d\hat{\psi}_j}{d\xi} \frac{2}{h^K} \right) \left( \frac{d\hat{\psi}_i}{d\xi} \frac{2}{h^K} \right) \frac{h^K}{2} d\xi \\ &= \frac{h^K}{2} \int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) d\xi \\ &\quad + \frac{2}{h^K} \int_{-1}^1 q \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \frac{d\hat{\psi}_j}{d\xi} \frac{d\hat{\psi}_i}{d\xi} d\xi \\ f_i^K &= \frac{h^K}{2} \int_{-1}^1 r \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_i(\xi) d\xi \\ &\quad - \frac{h^K}{2} \int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \left( \sum_{j=1}^{n_d^K} u_{g_j}^K \hat{\psi}_j(\xi) \right) \hat{\psi}_i(\xi) d\xi \\ &\quad - \frac{2}{h^K} \int_{-1}^1 q \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \left( \sum_{j=1}^{n_d^K} u_{g_j}^K \frac{d\hat{\psi}_j(\xi)}{d\xi} \right) \frac{d\hat{\psi}_i}{d\xi} d\xi \\ s_i^K &= s_{11}^K \hat{\psi}_i(-1) + s_{12}^K \hat{\psi}_i(1) \end{aligned}$$

### 5.1.5 Construction des fonctions d'interpolation $\hat{\psi}_i(\xi)$

À chaque noeud de calcul  $(x_1^K, x_2^K, \dots, x_{n_c^K}^K)$  de l'élément réel  $K$ , correspond un noeud d'interpolation  $(\xi_1, \xi_2, \dots, \xi_{n_c^K})$  sur l'élément de référence par la relation :

$$\xi_i = (T^K)^{-1}(x_i^K) \quad \text{ou encore} \quad x_i^K = T^K(\xi_i), \quad i = 1, 2, \dots, n_c^K$$

Ainsi, par construction, on aura :

$$\psi_j^K(x_i^K) = \hat{\psi}_j(\xi_i)$$

Puisque nous souhaitons calculer une solution approximative de l'équation différentielle d'ordre 2 de départ, il serait intéressant d'obtenir une approximation de  $u(x)$  à chaque noeud de calcul de chaque élément du domaine  $\Omega$ . De plus, puisque nous avons une équation différentielle d'ordre 2, la solution  $u(x)$  devra être dans  $H^1(\Omega)$ . L'idée de base est alors d'utiliser des approximations polynômiales sur chaque élément. L'approximation de la solution sera donc constituée de polynômes différents dans chaque élément. Pour s'assurer que cette approximation soit dans  $H^1(\Omega)$ , il faut s'assurer de la continuité de l'approximation à la frontière des éléments (voir le chapitre 2). Pour ce faire, il suffit d'imposer en chaque noeud de calcul de chaque élément :

$$u(x_i^K) = u^K(x_i^K) = \sum_{j=1}^{n_d^K} u_j^K \psi_j^K(x_i^K) = \sum_{j=1}^{n_d^K} u_j^K \hat{\psi}_j(\xi_i) = u_i^K \quad (5.10)$$

Ainsi, l'approximation calculée en  $x = x_i^K$  de  $u(x_i^K)$  sera  $u_i^K$  ce qui donne une interprétation physique des degrés de liberté  $u_i^K$ . Si un noeud de calcul est commun à plusieurs éléments, le degré de liberté associé à ce noeud sera toujours le même et la continuité de l'approximation sera assurée.

Pour satisfaire l'équation 5.10, il faut donc construire sur l'élément de référence les fonctions  $\hat{\psi}_j(\xi)$  de sorte que :

$$\hat{\psi}_j(\xi_i) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (5.11)$$

qui est la définition même des fonctions d'interpolation de Lagrange (voir l'annexe C ou Fortin, réf. [17]). Leur construction est donc immédiate. Il suffit maintenant de fixer le degré des polynômes que nous voulons utiliser dans chaque élément ce qui déterminera également la dimension  $n_d^K$  du système élémentaire.

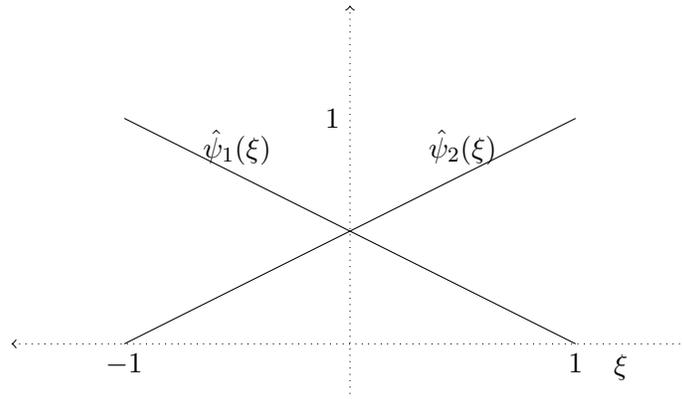
Remarquons enfin qu'avec des fonctions d'interpolation vérifiant l'équation 5.11 et puisque nous avons convenu que les 2 premiers noeuds correspondent aux extrémités de l'élément, le vecteur  $S^K$  des variables secondaires sur chaque élément est de la forme :

$$S^K = \begin{bmatrix} s_{11}^K \\ s_{12}^K \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

mettant ainsi en évidence que ces variables n'agissent qu'aux extrémités de l'élément.

### Approximation linéaire

Pour une approximation linéaire, il suffit de prendre 2 noeuds de calcul par élément et 1 seul degré de liberté par noeud de calcul ( $n_c^K = n_d^K = 2$ ). Les noeuds de calcul coïncident donc avec les

Figure 5.4 – Fonctions d’interpolation linéaires sur  $\hat{K}$ 

noeuds géométriques. Sur l’élément de référence les noeuds d’interpolation sont tout simplement les points  $\xi_1 = -1$  et  $\xi_2 = 1$ . Les fonctions  $\hat{\psi}_j(\xi)$  de Lagrange 5.11 sont :

$$\hat{\psi}_1(\xi) = \frac{(\xi - 1)}{(-1 - 1)} = \frac{(1 - \xi)}{2} \quad \text{et} \quad \hat{\psi}_2(\xi) = \frac{(\xi - (-1))}{(1 - (-1))} = \frac{(\xi + 1)}{2}$$

et sont illustrées à la figure 5.4. Pour évaluer le système élémentaire, on a besoin des dérivées :

$$\frac{d\hat{\psi}_1(\xi)}{d\xi} = -\frac{1}{2} \quad \text{et} \quad \frac{d\hat{\psi}_2(\xi)}{d\xi} = +\frac{1}{2}$$

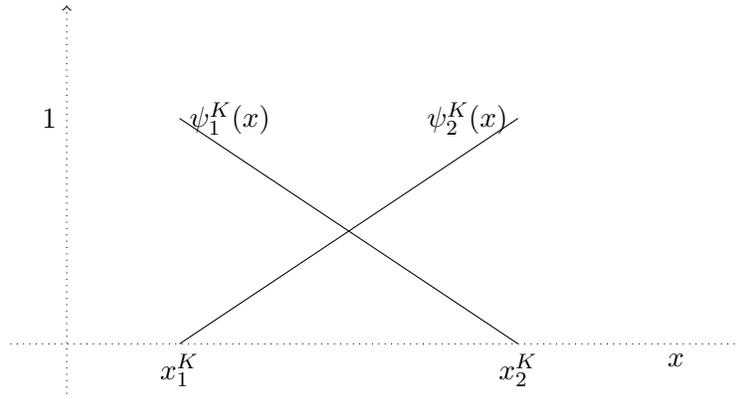
La formule de dérivation en chaîne nous donne alors :

$$\frac{d\psi_1^K(x)}{dx} = -\frac{1}{h^K} \quad \text{et} \quad \frac{d\psi_2^K(x)}{dx} = +\frac{1}{h^K}$$

Les fonctions d’interpolation linéaires sur l’élément de référence  $\hat{K}$  et sur l’élément  $K$  sont illustrées aux figures 5.4 et 5.5.

### Approximation quadratique

Pour une approximation quadratique, il faut 3 noeuds de calcul par élément (et encore un degré de liberté par noeud c.-à-d.  $n_c^K = n_d^K = 3$ ). On choisit d’abord les 2 extrémités de l’élément (les noeuds géométriques) et le troisième noeud est habituellement le point milieu  $\xi = 0$  de l’élément de référence, bien que ce ne soit absolument pas obligatoire. On a donc les noeuds d’interpolation

Figure 5.5 – Fonctions d’interpolation linéaires sur  $K$ 

$\xi_1 = -1$ ,  $\xi_2 = +1$  et  $\xi_3 = 0$ . Les fonctions  $\hat{\psi}_j(\xi)$  de Lagrange 5.11 de degré 2 sont alors :

$$\hat{\psi}_1(\xi) = \frac{(\xi - 0)(\xi - 1)}{(-1 - 0)(-1 - 1)} = \frac{\xi(\xi - 1)}{2}$$

$$\hat{\psi}_2(\xi) = \frac{(\xi - (-1))(\xi - 0)}{(1 - (-1))(1 - 0)} = \frac{\xi(\xi + 1)}{2}$$

$$\hat{\psi}_3(\xi) = \frac{(\xi - (-1))(\xi - 1)}{(0 - (-1))(0 - 1)} = 1 - \xi^2$$

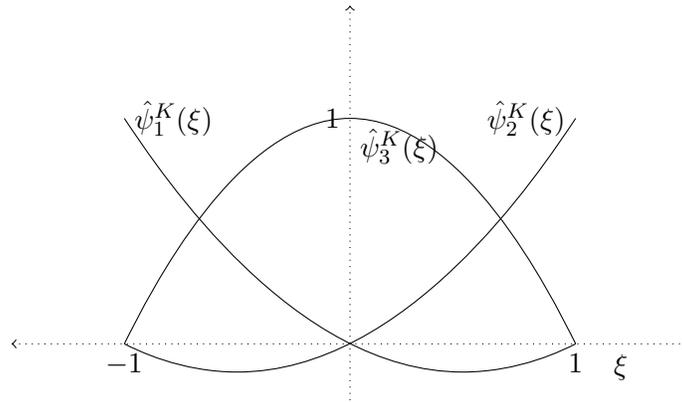
Ces fonctions sont illustrées à la figure 5.6 sur l’élément de référence et à la figure 5.7 sur l’élément réel. Pour évaluer le système élémentaire, on a encore ici besoin des dérivées que nous calculons une fois pour toutes :

$$\frac{d\hat{\psi}_1(\xi)}{d\xi} = \frac{(2\xi - 1)}{2}$$

$$\frac{d\hat{\psi}_2(\xi)}{d\xi} = \frac{(2\xi + 1)}{2}$$

$$\frac{d\hat{\psi}_3(\xi)}{d\xi} = -2\xi$$

La formule de dérivation en chaîne nous donne les dérivées des fonctions d’interpolation sur

Figure 5.6 – Fonctions d’interpolation quadratiques sur  $\hat{K}$ 

l’élément  $K$  :

$$\frac{d\psi_1^K(x)}{dx} = \frac{(2\xi - 1)}{h^K}$$

$$\frac{d\psi_2^K(x)}{dx} = \frac{(2\xi + 1)}{h^K}$$

$$\frac{d\psi_3^K(x)}{dx} = -\frac{4\xi}{h^K}$$

### Cas général

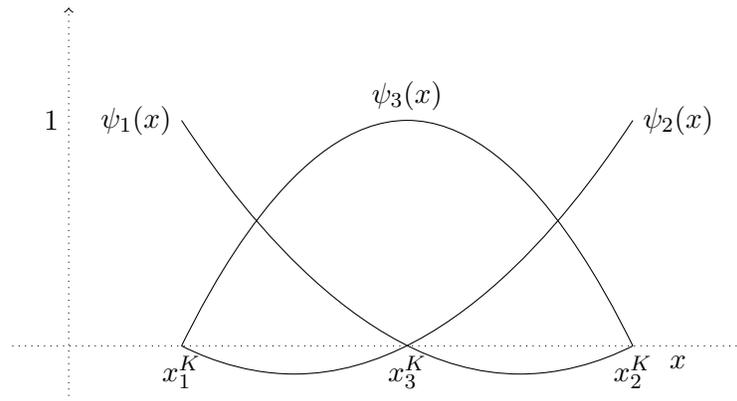
Il est maintenant facile de concevoir le cas général. Pour construire une approximation de degré quelconque  $m$ , il suffit de choisir sur l’élément de référence  $m + 1$  noeuds d’interpolation (incluant les extrémités de l’élément). On a ainsi  $n_c^K = m + 1$  noeuds de calcul et il suffit d’associer à chacun de ces noeuds une fonction d’interpolation de Lagrange  $\hat{\psi}_i(\xi)$  de degré  $m$ .

#### ◆ Remarque 5.3

Dans ce qui précède, nous avons utilisé des bases classiques (de Lagrange) de fonctions d’interpolation qui vérifient :

$$\hat{\psi}_j(\xi_i) = I_i^j$$

Cette contrainte n’est nullement obligatoire et on peut concevoir des bases différentes. C’est le cas par exemple des bases dites *hiérarchiques*. Ce concept mérite qu’on s’y attarde puisque le choix de la base est important, non seulement pour l’implantation de la méthode mais aussi pour des considérations numériques relatives au conditionnement de la matrice du système global. L’idée de base en dimension 1 consiste à utiliser des fonctions d’interpolation linéaires pour les 2 extrémités

Figure 5.7 – Fonctions d’interpolation quadratiques sur  $K$ 

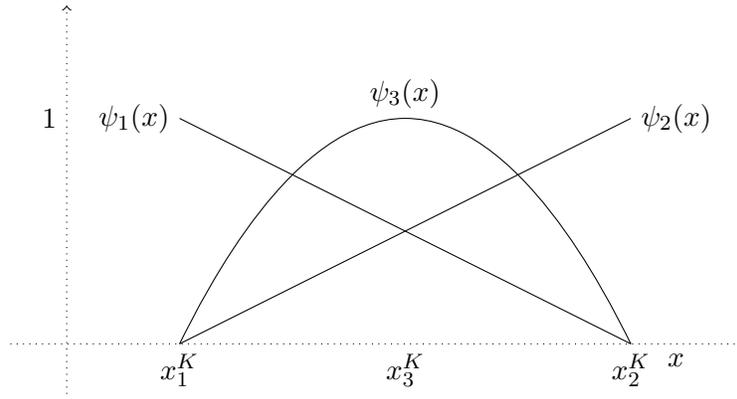
de l’élément et une fonction d’interpolation quadratique pour le point-milieu tel qu’illustré à la figure 5.8. Il résulte généralement de ces bases une interprétation moins évidente de la valeur des degrés de liberté associés aux noeuds de calcul. Nous reviendrons plus tard sur ces bases particulières. ♦

### 5.1.6 Évaluation du système élémentaire

Nous pouvons dès maintenant évaluer tous les coefficients du système élémentaire 5.5. En effet, toutes les fonctions nécessaires sont maintenant connues et il reste à effectuer les diverses intégrales. Deux options s’offrent à nous : on peut intégrer analytiquement ou recourir à l’intégration numérique.

#### Intégration analytique

C’est souvent la meilleure solution mais elle n’est pas toujours simple. Dans le cas de polynômes de bas degré et si les propriétés physiques (les fonctions  $p(x)$ ,  $q(x)$  et  $r(x)$ ) sont simples (par exemple constantes par élément), on choisira cette option. On obtient ainsi des expressions explicites pour les coefficients du système élémentaire. Ce travail peut devenir très pénible dans le cas général. Toutefois, on peut procéder de manière efficace en s’aidant de logiciels de calcul symbolique tels *Maple* ou *Mathematica*. Cette option est de plus en plus utilisée.

Figure 5.8 – Fonctions d'interpolation hiérarchiques sur  $K$ 

### Intégration numérique

C'est l'option la plus répandue et la plus versatile. Après le passage à l'élément de référence, les différents coefficients du système élémentaire requièrent l'évaluation d'intégrales de la forme :

$$I = \int_{-1}^1 g(\xi) d\xi \quad (5.12)$$

où la fonction  $g(\xi)$  fait intervenir les fonctions d'interpolation  $\hat{\psi}_i(\xi)$  et les propriétés physiques du problème. Par exemple, on doit évaluer :

$$\int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) \frac{h^K}{2} d\xi$$

qui est bien de la forme 5.12. Dans la plupart des programmes d'éléments finis, on utilise les quadratures de Gauss-Legendre qui consistent à approcher l'intégrale 5.12 par une expression de la forme :

$$\int_{-1}^1 g(\xi) d\xi \simeq \sum_{i=1}^{m_G} w_i g(\xi_i) \quad (5.13)$$

qui soit la plus précise possible. On présente un sommaire des techniques d'intégration numérique à l'annexe D. Le tableau D.1 résume quelques unes de ces quadratures en dimension 1. La dernière colonne de ce tableau fournit le degré des polynômes pour lesquels la quadrature de Gauss-Legendre est exacte et qui vaut  $2m_G - 1$ . C'est ce que l'on appelle le *degré de précision* de la formule de quadrature. En pratique, on choisit le nombre de points de Gauss-Legendre  $m_G$  en fonction des intégrales que l'on doit évaluer. Cela dépend donc du degré des fonctions  $\hat{\psi}_i(\xi)$  mais aussi des propriétés physiques. Par exemple, si la fonction  $p(x)$  est constante par élément ( $p(x) = p^K$  sur l'élément  $K$ ) et si des fonctions  $\hat{\psi}_i(\xi)$  quadratiques sont utilisées, alors on a :

$$\int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) \frac{h^K}{2} d\xi = \frac{p^K h^K}{2} \int_{-1}^1 \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) d\xi$$

La fonction à intégrer est de degré 4 (produit de 2 polynômes de degré 2) et une quadrature à 3 points ( $m_G = 3$ ) serait suffisante pour intégrer exactement. Par contre, des fonctions d'interpolation linéaires ne requéreraient qu'une quadrature de Gauss-Legendre à 2 points. Pour effectuer ce choix, il faut analyser toutes les intégrales apparaissant dans le système élémentaire et déterminer le degré de précision nécessaire.

Notons enfin que dans certaines situations, les quadratures de Gauss-Legendre ne seront jamais exactes. Par exemple, si  $p(x) = 1/x$ , la fonction à intégrer n'est plus polynomiale et les quadratures de Gauss-Legendre fournissent maintenant des approximations des coefficients du système élémentaire. Le choix du nombre de points de Gauss-Legendre  $m_G$  est alors plus délicat. Enfin, il convient de souligner que le coût total de l'assemblage (voir la section 5.1.7) est proportionnel au nombre de points de Gauss-Legendre utilisés et que le choix de la quadrature doit aussi tenir compte de cette contrainte.

### Cas particuliers

Puisque dans plusieurs exemples, les fonctions  $p(x)$ ,  $q(x)$  et  $r(x)$  sont constantes par élément (et parfois même constantes dans tout le domaine), il est utile de calculer immédiatement et une fois pour toutes le système élémentaire dans ce cas particulier. On notera  $p^K$ ,  $q^K$  et  $r^K$  les valeurs respectives de ces fonctions sur l'élément  $K$ . Nous supposons de plus que le relèvement  $u_g(x)$  des conditions aux limites essentielles est nul. Les différentes contributions au système élémentaire sont alors :

$$\begin{aligned} a_{ij}^K &= \frac{h^K p^K}{2} \int_{-1}^1 \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) d\xi + \frac{2q^K}{h^K} \int_{-1}^1 \frac{d\hat{\psi}_j}{d\xi} \frac{d\hat{\psi}_i}{d\xi} d\xi \\ f_i^K &= \frac{h^K}{2} \int_{-1}^1 r \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_i(\xi) d\xi = \frac{h^K r^K}{2} \int_{-1}^1 \hat{\psi}_i(\xi) d\xi \\ s_i^K &= s_{11}^K \hat{\psi}_i(-1) + s_{12}^K \hat{\psi}_i(1) \end{aligned}$$

#### – Interpolation linéaire

Si on utilise des fonctions d'interpolation linéaires, on remarque que les coefficients de la matrice élémentaire sont très faciles à calculer puisque l'intégrand est alors un polynôme de degré maximum 2. On peut les évaluer analytiquement en notant toutefois au passage qu'une quadrature de Gauss-Legendre à 2 points donnerait le même résultat. On obtient ainsi :

$$A^K = \frac{h^K p^K}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \frac{q^K}{h^K} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad (5.14)$$

Quant au membre de droite, l'intégrant est linéaire et une formule de Gauss-Legendre à 1 point suffirait. On a :

$$f^K + s^K = \frac{h^K r^K}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} s_{11}^K \\ s_{12}^K \end{bmatrix} \quad (5.15)$$

– **Interpolation quadratique**

Si on utilise des fonctions d'interpolation quadratiques, les coefficients de la matrice élémentaire résultent en des intégrants de degré maximum 4. On peut encore les évaluer analytiquement mais une quadrature de Gauss-Legendre à 3 points donnerait le même résultat. On obtient ainsi :

$$A^K = \frac{h^K p^K}{30} \begin{bmatrix} 4 & -1 & 2 \\ -1 & 4 & 2 \\ 2 & 2 & 16 \end{bmatrix} + \frac{q^K}{3h^K} \begin{bmatrix} 7 & 1 & -8 \\ 1 & 7 & -8 \\ -8 & -8 & 16 \end{bmatrix} \quad (5.16)$$

En ce qui concerne le membre de droite, l'intégrant est quadratique et une formule de Gauss-Legendre à 2 points suffirait. On a :

$$f^K + s^K = \frac{h^K r^K}{6} \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} + \begin{bmatrix} s_{11}^K \\ s_{12}^K \\ 0 \end{bmatrix} \quad (5.17)$$

Ces systèmes élémentaires reviennent fréquemment dans les exemples et les exercices de telle sorte qu'on y référera aux moments opportuns. Il importe cependant de bien comprendre comment on les obtient. Notons enfin que dans le cas où les fonctions  $p(x)$ ,  $q(x)$  et  $r(x)$  ne sont plus constantes par élément, les équations 5.14 à 5.17 ne sont plus valides.

### 5.1.7 Assemblage

L'étape de l'assemblage consiste à prendre en compte les contributions de tous les systèmes élémentaires pour construire un système linéaire global que l'on devra résoudre tout comme on l'a fait pour la méthode de Ritz. La clé de l'assemblage est le tableau d'adressage des degrés de liberté *adres* qui permet de passer du système élémentaire local (sur un élément  $K$ ) au système global (sur tout le domaine  $\Omega$ ) en fonction de la numérotation des degrés de liberté. Avant de procéder à l'assemblage comme tel, nous établissons le lien entre la méthode de Ritz et celle des éléments finis en montrant comment cette méthode permet une construction automatique des fonctions de Ritz.

#### Construction des fonctions de Ritz

On associe une fonction de Ritz à chaque degré de liberté du domaine. C'est donc le tableau *adres* qui permet de construire les fonctions de Ritz sur tout le domaine  $\Omega$ . Pour ce faire, regardons ce qui se produit pour chacun des *nddl* degrés de liberté du domaine. Deux situations peuvent survenir suivant que le degré de liberté est commun ou non à plusieurs éléments.

La situation la plus simple est celle où le degré de liberté  $i$  étudié ( $0 \leq i \leq nddl$ ) n'appartient qu'à un seul élément  $K$ . Le numéro  $i$  n'apparaît donc qu'une seule fois dans le tableau d'adressage à la ligne  $K$  ( $K$  désigne à la fois l'élément lui-même et son numéro). Le tableau *adres* nous indique à quel degré de liberté de l'élément  $K$ . Supposons donc que ce  $i^{\text{ième}}$  degré de liberté du domaine soit le  $k^{\text{ième}}$  degré de liberté ( $1 \leq k \leq n_d^K$ ) de l'élément  $K$  c.-à-d.

$$adres(K, k) = i$$

Sur  $K$ , ce degré de liberté est associé à une fonction d'interpolation  $\psi_k^K(x)$ . Cette fonction d'interpolation n'est définie que dans l'élément  $K$  et peut être prolongée par 0 à l'extérieur de  $K$ . La fonction ainsi obtenue est la fonction de Ritz associée à ce degré de liberté qui correspond globalement à la  $i^e$  fonction de Ritz du domaine. Notons immédiatement que le *support* de cette fonction de Ritz se réduit au seul élément  $K$  et on a :

$$\phi_i(x) = \begin{cases} \psi_k^K(x) & \text{si } x \in K \\ 0 & \text{ailleurs} \end{cases}$$

Supposons maintenant que le  $j^{\text{ième}}$  degré de liberté (pour un certain  $j$  compris entre 1 et  $nddl$ ) du domaine soit commun à 2 éléments. Un degré de liberté peut être commun à plus de 2 éléments et le raisonnement qui suit se généralise facilement. La fonction de Ritz associée est constituée de 2 parties. Notons  $K_1$  et  $K_2$  les 2 éléments en question. Le numéro  $j$  apparaît donc uniquement aux lignes  $K_1$  et  $K_2$  du tableau d'adressage. Sur le premier élément, le degré de liberté  $j$  peut correspondre au  $k_1^{\text{ième}}$  degré de liberté de l'élément  $K_1$  et au  $k_2^{\text{ième}}$  degré de liberté de l'élément  $K_2$  ( $1 \leq k_1, k_2 \leq n_d^K$ ). On a donc  $adres(K_1, k_1) = adres(K_2, k_2) = j$ . La fonction de Ritz associée sera donc définie par :

$$\phi_j(x) = \begin{cases} \psi_{k_1}^{K_1}(x) & \text{si } x \in K_1 \\ \psi_{k_2}^{K_2}(x) & \text{si } x \in K_2 \\ 0 & \text{ailleurs} \end{cases}$$

Notons en terminant que le support de cette fonction de Ritz se réduit aux 2 éléments  $K_1$  et  $K_2$ .

► **Exemple 5.2**

Considérons un maillage de 3 éléments tel qu'illustré à la figure 5.9. Nous supposons de plus que les fonctions d'interpolation sont linéaires et qu'il n'y a qu'un seul degré de liberté par noeud de calcul ( $n_c^K = n_g^K = n_d^K = 2$ ). Pour simplifier la présentation, nous supposons que :

$$numer(j) = j \quad j = 1, 2, 3, \dots, nddl$$

de sorte que les tableaux *connec* et *adres* coïncident, ce qui revient à dire qu'aucune variable essentielle n'est imposée. Il en résulte le tableau d'adressage :

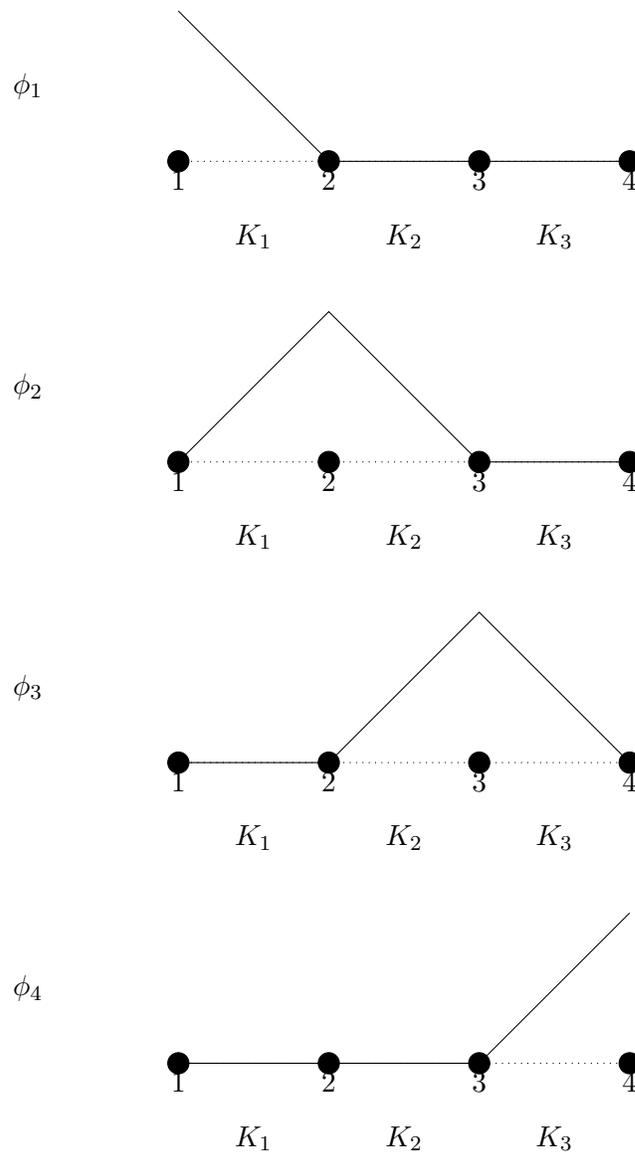


Figure 5.9 – Fonctions de Ritz linéaires par élément

Numéros des ddl		
Tableau <i>adres</i>		
Élément	Ddl #1	Ddl #2
1	1	2
2	2	3
3	3	4

Il y a donc au total 4 degrés de liberté ( $nddl = 4$ ) et la figure 5.9 présente les 4 fonctions de Ritz qui y sont associées. Les fonctions  $\phi_1(x)$  et  $\phi_4(x)$  ne sont non nulles que respectivement sur les éléments  $K_1$  et  $K_3$ . Partout ailleurs, elles sont nulles. On remarque de plus que ces numéros (1 et 4) n'apparaissent respectivement dans le tableau d'adressage qu'aux lignes 1 et 3 respectivement, correspondant aux éléments 1 et 3. Puisque  $adres(1, 1) = 1$  et  $adres(3, 2) = 4$ , on a :

$$\phi_1(x) = \begin{cases} \psi_1^{K_1}(x) & \text{si } x \in K_1 \\ 0 & \text{ailleurs} \end{cases}$$

et :

$$\phi_4(x) = \begin{cases} \psi_2^{K_3}(x) & \text{si } x \in K_3 \\ 0 & \text{ailleurs} \end{cases}$$

Par contre, la fonction  $\phi_2(x)$  est à cheval sur les éléments 1 et 2. On remarque alors que  $adres(1, 2) = adres(2, 1) = 2$ . On a dans ce cas :

$$\phi_2(x) = \begin{cases} \psi_2^{K_1}(x) & \text{si } x \in K_1 \\ \psi_1^{K_2}(x) & \text{si } x \in K_2 \\ 0 & \text{ailleurs} \end{cases}$$

De même, puisque  $adres(2, 2) = adres(3, 1) = 3$

$$\phi_3(x) = \begin{cases} \psi_2^{K_2}(x) & \text{si } x \in K_2 \\ \psi_1^{K_3}(x) & \text{si } x \in K_3 \\ 0 & \text{ailleurs} \end{cases}$$



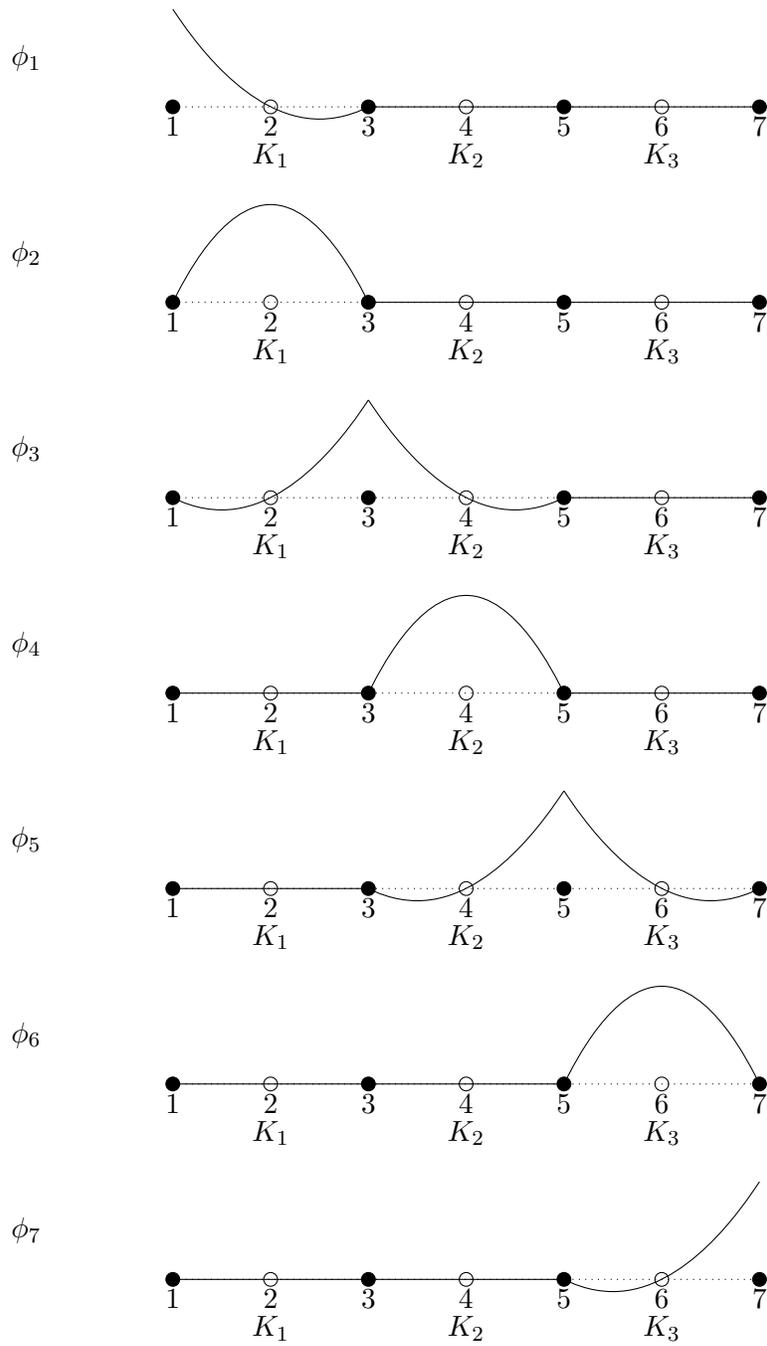


Figure 5.10 – Fonctions de Ritz quadratiques par élément

► **Exemple 5.3**

Sur le même maillage de 3 éléments de la figure 5.9 et en faisant les mêmes hypothèses concernant la numérotation des degrés de liberté, on considère cette fois des fonctions d'interpolation quadratiques ( $n_c^K = n_d^K = 3, n_g^K = 2$ ). Ici encore, on a numéroté les degrés de liberté du domaine tout simplement de gauche à droite. Les tableaux de connectivité *connec* et *adres* sont alors les mêmes :

Numéros des ddl			
Tableau <i>adres</i>			
Élément	Ddl #1	Ddl #2	Ddl #3
1	1	3	2
2	3	5	4
3	5	7	6

Il y a donc au total 7 degrés de liberté ( $n_{ddl} = 7$ ) et la figure 5.10 présente les 7 fonctions de Ritz qui y sont associées. On remarque que ces fonctions sont maintenant quadratiques (au lieu de linéaires) sur les éléments qui constituent leur support. Partout ailleurs, elles sont nulles. Par exemple, puisque  $adres(1, 3) = 2$ , on a :

$$\phi_2(x) = \begin{cases} \psi_3^{K_1}(x) & \text{si } x \in K_1 \\ 0 & \text{ailleurs} \end{cases}$$

et puisque  $adres(2, 2) = adres(3, 1) = 5$  :

$$\phi_5(x) = \begin{cases} \psi_2^{K_2}(x) & \text{si } x \in K_2 \\ \psi_1^{K_3}(x) & \text{si } x \in K_3 \\ 0 & \text{ailleurs} \end{cases}$$

Il en va de manière similaire pour les autres fonctions de Ritz. ◀

◆ **Remarque 5.4**

Notons enfin, et cela est fondamental, que le support des fonctions de Ritz se réduit à très peu d'éléments. Ainsi, pour obtenir les coefficients du système linéaire global de la méthode de Ritz  $a(\phi_j, \phi_i)$ , il n'est pas nécessaire d'intégrer sur tout le domaine puisque le support des fonctions  $\phi_i(x)$  est très réduit. De plus, l'intersection des supports des fonctions  $\phi_i(x)$  et  $\phi_j(x)$  est parfois nul ce qui entraîne que  $a(\phi_j, \phi_i) = 0$ . L'intégration ne porte donc que sur l'intersection des supports des fonctions  $\phi_i(x)$  et  $\phi_j(x)$ . Il en résulte une structure matricielle très particulière pour le système linéaire global. On parle alors de *matrice creuse* signifiant ainsi que la matrice obtenue contient une

part importante de zéros. On profitera éventuellement de cette structure pour diminuer l'espace-mémoire requis en ne conservant que les termes non nuls. Nous reviendrons sur cette question un peu plus loin. ♦

Mais comment s'y retrouver pour éviter des calculs inutiles ? C'est encore le tableau d'adressage qui nous guide. Il suffit en effet de calculer les systèmes élémentaires sur chaque élément et le tableau d'adressage nous indique à quel endroit apporter cette contribution dans le système global.

### Construction du système global

À partir des systèmes élémentaires sur chaque élément  $K$ , on construit un système d'équations linéaires global qui tient compte des contributions de chaque degré de liberté de chaque élément. La matrice globale sera notée  $A$  (et ses coefficients  $a_{ij}$ ). Rappelons que le vecteur global des degrés de liberté est noté  $U$  (et ses coefficients  $u_i$ ). Le terme de droite sera encore constitué de 2 parties  $F$  et  $S$  (de coefficients respectifs  $f_i$  et  $s_i$ ). Pour saisir comment s'y prendre, il suffit de remarquer que pour notre problème type :

$$\begin{aligned} a_{ij} &= a(\phi_j(x), \phi_i(x)) = \int_0^1 \left( p(x)\phi_j(x)\phi_i(x) + q(x)\frac{d\phi_j}{dx}\frac{d\phi_i}{dx} \right) dx \\ &= \sum_{K=1}^{nel} \int_K \left( p(x)\phi_j(x)|_K \phi_i(x)|_K + q(x)\frac{d\phi_j}{dx}\Big|_K \frac{d\phi_i}{dx}\Big|_K \right) dx \end{aligned}$$

Comme nous l'avons déjà souligné, la dernière somme ne porte que sur les éléments faisant partie du support à la fois de la fonction  $\phi_i(x)$  et de la fonction  $\phi_j(x)$ . Sur le plan pratique, cela signifie que les numéros  $i$  et  $j$  apparaissent tous les deux à la ligne du tableau *adres* correspondant à ces éléments. Soit donc  $K^*$  l'un de ces éléments. Les degrés de liberté  $i$  et  $j$  apparaissent forcément dans le vecteur d'adressage de cet élément (la  $K^*$  ligne du tableau *adres*). On aura par exemple que  $adres(K^*, k_1) = i$  et  $adres(K^*, k_2) = j$  où  $k_1$  et  $k_2$  sont des nombres compris entre 1 et  $n_d^K$ . On a alors :

$$\begin{aligned} &\int_{K^*} \left( p(x)\phi_j(x)|_{K^*} \phi_i(x)|_{K^*} + q(x)\frac{d\phi_j}{dx}\Big|_{K^*} \frac{d\phi_i}{dx}\Big|_{K^*} \right) dx \\ &= \int_{K^*} \left( p(x)\psi_{k_2}^{K^*}(x)\psi_{k_1}^{K^*}(x) + q(x)\frac{d\psi_{k_2}^{K^*}}{dx}\frac{d\psi_{k_1}^{K^*}}{dx} \right) dx \\ &= a_{k_1 k_2}^{K^*} \end{aligned}$$

On ajoute ce dernier coefficient au terme  $a_{ij}$  du système global et on fera de même pour tous les éléments dont le vecteur d'adressage contient à la fois les numéros  $i$  et  $j$ . On constate donc que chaque coefficient de la matrice élémentaire apporte une contribution au système linéaire global et qu'il suffit de sommer les contributions de chaque système élémentaire.

### Assemblage des matrices élémentaires

En pratique, la procédure d'assemblage requiert de suivre le chemin inverse de celui que nous venons de décrire et de partir du système élémentaire sur chaque élément  $K$  :

$$A^K \delta_U^K = F^K + S^K$$

ou plus explicitement :

$$\begin{bmatrix} a_{11}^K & a_{12}^K & \cdots & a_{1n_d}^K \\ a_{21}^K & a_{22}^K & \cdots & a_{2n_d}^K \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_d 1}^K & a_{n_d 2}^K & \cdots & a_{n_d n_d}^K \end{bmatrix} \begin{bmatrix} \delta_{u_1}^K \\ \delta_{u_2}^K \\ \vdots \\ \delta_{u_{n_d}}^K \end{bmatrix} = \begin{bmatrix} f_1^K \\ f_2^K \\ \vdots \\ f_{n_d}^K \end{bmatrix} + \begin{bmatrix} s_{11}^K \\ s_{12}^K \\ \vdots \\ 0 \end{bmatrix}$$

Considérons l'élément  $a_{k_1 k_2}^K$  de ce système élémentaire. Le tableau d'adressage nous indique par exemple que  $adres(K, k_1) = i$  et que  $adres(K, k_2) = j$ . On a alors :

$$\begin{aligned} a_{k_1 k_2}^K &= \int_K \left( p(x) \psi_{k_2}^K(x) \psi_{k_1}^K(x) + q(x) \frac{d\psi_{k_2}^K}{dx} \frac{d\psi_{k_1}^K}{dx} \right) dx \\ &= \int_K \left( p(x) \phi_j(x)|_K \phi_i(x)|_K + q(x) \frac{d\phi_j}{dx} \Big|_K \frac{d\phi_i}{dx} \Big|_K \right) dx \end{aligned}$$

et que cette contribution doit être ajoutée au coefficient  $a_{ij}$  du système global.

$$A \delta_U = F + S$$

De même, on ajoute les coefficients locaux  $f_{k_1}^K$  et  $s_{k_1}^K$  aux coefficients globaux  $f_i$  et  $s_i$ .

De toute la discussion qui précède, on conclut que l'on construit la matrice  $A$  et les vecteurs  $F$  et  $S$  de la manière suivante, qui constitue l'assemblage :

- Initialisation à 0 de la matrice  $A$  et des vecteurs  $F$  et  $S$  ;
- Pour chaque élément  $K$  ;
  - Pour chaque degré de liberté  $k_1 = 1, 2, \dots, n_d^K$  ;
    - Numéro de la ligne :  $i = adres(K, k_1)$
    - $f_i \leftarrow f_i + f_{k_1}^K$
    - $s_i \leftarrow s_i + s_{k_1}^K$
  - Pour chaque degré de liberté  $k_2 = 1, 2, \dots, n_d^K$  ;
    - Numéro de la colonne :  $j = adres(K, k_2)$
    - $a_{ij} \leftarrow a_{ij} + a_{k_1 k_2}^K$
  - Fin de la boucle sur les colonnes
- Fin de la boucle sur les lignes
- Fin de la boucle sur les éléments

◆ **Remarque 5.5**

Il est concevable de construire des tableaux d'adressage différents pour les lignes et les colonnes d'une matrice. On rencontre cette situation notamment dans les problèmes ayant plusieurs inconnues. On peut ainsi assembler des matrices rectangulaires si nécessaire. ◆

► **Exemple 5.4**

Considérons une fois encore le maillage très simple de 3 éléments quadratiques de la figure 5.3. On trouve le tableau d'adressage à la page 84. Le système global sera de dimension 7 sur 7. Pour le construire, on va assembler les 3 systèmes élémentaires de dimension 3 de la forme :

$$\begin{bmatrix} a_{11}^K & a_{12}^K & a_{13}^K \\ a_{21}^K & a_{22}^K & a_{23}^K \\ a_{31}^K & a_{32}^K & a_{33}^K \end{bmatrix} \begin{bmatrix} \delta_{u_1}^K \\ \delta_{u_2}^K \\ \delta_{u_3}^K \end{bmatrix} = \begin{bmatrix} f_1^K \\ f_2^K \\ f_3^K \end{bmatrix} + \begin{bmatrix} s_{11}^K \\ s_{12}^K \\ 0 \end{bmatrix}$$

Le vecteur d'adressage du premier élément  $K_1$  est  $[7, 2, 1]$ . Cela signifie que les inconnues élémentaires  $u_1^{K_1}$ ,  $u_2^{K_1}$  et  $u_3^{K_1}$  correspondent aux degrés de liberté  $u_7$ ,  $u_2$  et  $u_1$  du système global. Le système élémentaire est donc équivalent à :

$$\begin{bmatrix} a_{33}^{K_1} & a_{32}^{K_1} & 0 & 0 & 0 & 0 & a_{31}^{K_1} \\ a_{23}^{K_1} & a_{22}^{K_1} & 0 & 0 & 0 & 0 & a_{21}^{K_1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{13}^{K_1} & a_{12}^{K_1} & 0 & 0 & 0 & 0 & a_{11}^{K_1} \end{bmatrix} \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ \delta_{u_7} \end{bmatrix} = \begin{bmatrix} f_3^{K_1} \\ f_2^{K_1} \\ 0 \\ 0 \\ 0 \\ 0 \\ f_1^{K_1} \end{bmatrix} + \begin{bmatrix} 0 \\ s_{12}^{K_1} \\ 0 \\ 0 \\ 0 \\ 0 \\ s_{11}^{K_1} \end{bmatrix}$$

qui n'est qu'une réécriture différente du système élémentaire. Une façon systématique de s'y retrouver consiste à écrire le système élémentaire sous la forme :

$$\left[ \begin{array}{c|ccc} & 7 & 2 & 1 \\ \hline 7 & a_{11}^{K_1} & a_{12}^{K_1} & a_{13}^{K_1} \\ 2 & a_{21}^{K_1} & a_{22}^{K_1} & a_{23}^{K_1} \\ 1 & a_{31}^{K_1} & a_{32}^{K_1} & a_{33}^{K_1} \end{array} \right] \begin{bmatrix} \delta_{u_1}^{K_1} \\ \delta_{u_2}^{K_1} \\ \delta_{u_3}^{K_1} \end{bmatrix} = \begin{bmatrix} f_1^{K_1} \\ f_2^{K_1} \\ f_3^{K_1} \end{bmatrix} + \begin{bmatrix} s_{11}^{K_1} \\ s_{12}^{K_1} \\ 0 \end{bmatrix}$$

où on a simplement ajouté devant les lignes et au dessus des colonnes le vecteur d'adressage du premier élément (la première ligne du tableau *adres*). Pour connaître l'adresse dans le système global d'un coefficient du système élémentaire, il suffit de regarder le vecteur d'adressage de la ligne et/ou de la colonne. Ainsi, le coefficient  $a_{21}^{K_1}$  du système élémentaire apportera sa contribution au coefficient  $a_{27}$  du système global. De même, les coefficients  $f_1^{K_1}$  et  $s_{11}^{K_1}$  seront ajoutés aux coefficients  $f_7$  et  $s_7$  du vecteur des sollicitations global.

Pour le deuxième élément dont le vecteur d'adressage est  $[2, 4, 3]$ , un raisonnement similaire nous permet d'écrire :

$$\begin{bmatrix} & 2 & 4 & 3 \\ 2 & a_{11}^{K_2} & a_{12}^{K_2} & a_{13}^{K_2} \\ 4 & a_{21}^{K_2} & a_{22}^{K_2} & a_{23}^{K_2} \\ 3 & a_{31}^{K_2} & a_{32}^{K_2} & a_{33}^{K_2} \end{bmatrix} \begin{bmatrix} \delta_{u_1}^{K_2} \\ \delta_{u_2}^{K_2} \\ \delta_{u_3}^{K_2} \end{bmatrix} = \begin{bmatrix} f_1^{K_2} \\ f_2^{K_2} \\ f_3^{K_2} \end{bmatrix} + \begin{bmatrix} s_{11}^{K_2} \\ s_{12}^{K_2} \\ 0 \end{bmatrix}$$

qui devient :

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{11}^{K_2} & a_{13}^{K_2} & a_{12}^{K_2} & 0 & 0 & 0 \\ 0 & a_{31}^{K_2} & a_{33}^{K_2} & a_{32}^{K_2} & 0 & 0 & 0 \\ 0 & a_{21}^{K_2} & a_{23}^{K_2} & a_{22}^{K_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ \delta_{u_7} \end{bmatrix} = \begin{bmatrix} 0 \\ f_1^{K_2} \\ f_3^{K_2} \\ f_2^{K_2} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ s_{11}^{K_2} \\ 0 \\ s_{12}^{K_2} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Pour le dernier élément, dont le vecteur d'adressage est [4, 6, 5], on obtient :

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{11}^{K_3} & a_{13}^{K_3} & a_{12}^{K_3} & 0 \\ 0 & 0 & 0 & a_{31}^{K_3} & a_{33}^{K_3} & a_{32}^{K_3} & 0 \\ 0 & 0 & 0 & a_{21}^{K_3} & a_{23}^{K_3} & a_{22}^{K_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ \delta_{u_7} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ f_1^{K_3} \\ f_3^{K_3} \\ f_2^{K_3} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ s_{11}^{K_3} \\ 0 \\ s_{12}^{K_3} \\ 0 \end{bmatrix}$$

Le système global est alors obtenu en sommant toutes les contributions :

$$\begin{bmatrix} a_{33}^{K_1} & a_{32}^{K_1} & 0 & 0 & 0 & 0 & a_{31}^{K_1} \\ a_{23}^{K_1} & a_{22}^{K_1} + a_{11}^{K_2} & a_{13}^{K_2} & a_{12}^{K_2} & 0 & 0 & a_{21}^{K_1} \\ 0 & a_{31}^{K_2} & a_{33}^{K_2} & a_{32}^{K_2} & 0 & 0 & 0 \\ 0 & a_{21}^{K_2} & a_{23}^{K_2} & a_{22}^{K_2} + a_{11}^{K_3} & a_{13}^{K_3} & a_{12}^{K_3} & 0 \\ 0 & 0 & 0 & a_{31}^{K_3} & a_{33}^{K_3} & a_{32}^{K_3} & 0 \\ 0 & 0 & 0 & a_{21}^{K_3} & a_{23}^{K_3} & a_{22}^{K_3} & 0 \\ a_{13}^{K_1} & a_{12}^{K_1} & 0 & 0 & 0 & 0 & a_{11}^{K_1} \end{bmatrix} \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ \delta_{u_7} \end{bmatrix} = \begin{bmatrix} f_3^{K_1} \\ f_2^{K_1} + f_1^{K_2} \\ f_3^{K_2} \\ f_2^{K_2} + f_1^{K_3} \\ f_3^{K_3} \\ f_2^{K_3} \\ f_1^{K_1} \end{bmatrix} + \begin{bmatrix} 0 \\ s_{12}^{K_1} + s_{11}^{K_2} \\ 0 \\ s_{12}^{K_2} + s_{11}^{K_3} \\ 0 \\ s_{12}^{K_3} \\ s_{11}^{K_1} \end{bmatrix} \quad (5.18)$$



◆ **Remarque 5.6**

Il est évident que seul ce dernier système de dimension  $nddl$  sur  $nddl$  est explicitement construit. On ne construit pas un tel système pour chaque élément pour ensuite les additionner comme nous l'avons fait. Cela nécessiterait une quantité de mémoire phénoménale et nous ne l'avons fait que pour illustrer le processus d'assemblage. On additionne les contributions de chaque système élémentaire directement dans le système 5.18 à l'adresse fournie par le vecteur d'adressage de chaque élément.



### 5.1.8 Imposition des conditions aux limites

On impose les conditions en deux étapes suivant que l'on traite les conditions essentielles (variables primaires) ou les conditions naturelles (variables secondaires). Une fois le système global assemblé, il prend la forme suivante :

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \delta_U^I \\ 0 \end{pmatrix} = \begin{pmatrix} F_1^C \\ F_2^C \end{pmatrix} + \begin{pmatrix} S^C \\ S^I \end{pmatrix} \quad (5.19)$$

Cette partition particulière provient de la numérotation des degrés de liberté que nous avons établie et aussi de la formulation en correction qui annule la dernière partie du vecteur  $\delta_U$ . La partition de la matrice  $A$  suit directement celle du vecteur global des degrés de liberté  $U$ . On note au passage que les matrices  $M_{11}$  et  $M_{22}$  sont carrées et que les matrices  $M_{12}$  et  $M_{21}$  sont rectangulaires. Si la forme bilinéaire du problème est symétrique, on a  $M_{12}^T = M_{21}$ . Nous reviendrons sur cette partition lors de la résolution du système global.

Enfin, il reste à analyser le terme de droite composé de 2 parties. Le vecteur  $F$  est entièrement déterminé et ne pose aucun problème. Par contre, le vecteur  $S$  contenant la contribution des variables secondaires est lui aussi décomposé en 2 parties. Là où la variable essentielle est imposée (et donc connue), nous avons vu que la condition naturelle est inconnue et vice versa. La situation est donc claire aux 2 extrémités du domaine.

Il reste à regarder ce qui se passe aux frontières entre les éléments. Typiquement, si le degré de liberté  $i$  du domaine est attaché à un noeud  $x$  à la frontière entre 2 éléments  $K_-$  et  $K_+$ , le vecteur  $S$  contiendra à la  $i^{\text{ème}}$  ligne une expression de la forme :

$$s_i = s_{12}^{K_-} + s_{11}^{K_+}$$

De la définition même de ces variables, on a :

$$s_i = q(x^-) \frac{du}{dx}(x^-) - q(x^+) \frac{du}{dx}(x^+)$$

ce qui signifie que  $s_i$  est le saut de la variable naturelle (secondaire) :

$$-q(x) \frac{du}{dx}(x)$$

au noeud  $x$ . Les indices  $-$  et  $+$  réfèrent aux valeurs à gauche (prise dans l'élément  $K_-$ ) et à droite (prise dans l'élément  $K_+$ ) de la variable en  $x$ . Or si ce saut était différent de 0, le terme de droite de l'équation différentielle 5.1 ferait apparaître une distribution de Dirac  $H\delta_x$ . On utilise fréquemment les distributions de Dirac pour modéliser une charge ponctuelle d'intensité  $H$ . Si tel est le cas, c'est au moment de l'imposition des conditions naturelles à ce noeud que l'on prend en compte la contribution de cette charge ponctuelle et on pose :

$$s_i = s_{12}^{K_-} + s_{11}^{K_+} = H$$

Sinon, on pose simplement  $s_i = 0$ .

### 5.1.9 Solution du système global

Pour la résolution du système linéaire, 2 autres étapes sont nécessaires. Tout d'abord, on détermine le vecteur  $U^I$  en résolvant le système linéaire :

$$M_{11}\delta_U^I = F_1^C + S_1^C \quad (5.20)$$

qui n'est qu'une réécriture de la première équation du système 5.19. On remarque que le terme de droite est entièrement connu. Pour cela, on utilise les techniques classiques comme la méthode de décomposition  $LU$ . Remarquons que cette équation n'est rien d'autre que la discrétisation de la forme variationnelle :

$$a(\delta_u, w) = l(w) - a(u_g, w)$$

Une fois le vecteur  $U^I$  calculé, on détermine si nécessaire le vecteur  $S^I$  directement en posant :

$$S^I = M_{21}\delta_U^I - F_2^C \quad (5.21)$$

On remarque enfin que les sous-matrices  $M_{12}$  et  $M_{22}$  ne jouent aucun rôle et pourraient tout simplement ne pas être assemblées.

#### ► Exemple 5.5

Nous sommes en mesure de compléter l'exemple illustré à la figure 5.3 qui nous a mené au système linéaire 5.18. Imposons d'abord les conditions aux limites. D'une part, le seul degré de liberté qui est fixé est le 7<sup>ième</sup>. À cet endroit,  $u(0) = c$ , ce qui entraîne que  $u_7 = c$  (et donc  $\delta_{u_7} = 0$ ). On peut donc partitionner le système 5.18 sous la forme :

$$\begin{bmatrix} a_{33}^{K_1} & a_{32}^{K_1} & 0 & 0 & 0 & 0 & a_{31}^{K_1} \\ a_{23}^{K_1} & a_{22}^{K_1} + a_{11}^{K_2} & a_{13}^{K_2} & a_{12}^{K_2} & 0 & 0 & a_{21}^{K_1} \\ 0 & a_{31}^{K_2} & a_{33}^{K_2} & a_{32}^{K_2} & 0 & 0 & 0 \\ 0 & a_{21}^{K_2} & a_{23}^{K_2} & a_{22}^{K_2} + a_{11}^{K_3} & a_{13}^{K_3} & a_{12}^{K_3} & 0 \\ 0 & 0 & 0 & a_{31}^{K_3} & a_{33}^{K_3} & a_{32}^{K_3} & 0 \\ 0 & 0 & 0 & a_{21}^{K_3} & a_{23}^{K_3} & a_{22}^{K_3} & 0 \\ \hline a_{13}^{K_1} & a_{12}^{K_1} & 0 & 0 & 0 & 0 & a_{11}^{K_1} \end{bmatrix} \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} f_3^{K_1} \\ f_2^{K_1} + f_1^{K_2} \\ f_3^{K_2} \\ f_2^{K_2} + f_1^{K_3} \\ f_3^{K_3} \\ f_2^{K_3} \\ \hline f_1^{K_1} \end{bmatrix} + \begin{bmatrix} 0 \\ s_{12}^{K_1} + s_{11}^{K_2} \\ 0 \\ s_{12}^{K_2} + s_{11}^{K_3} \\ 0 \\ s_{12}^{K_3} \\ \hline s_{11}^{K_1} \end{bmatrix}$$

où nous avons explicité la forme 5.19. On peut dès lors visualiser les matrices  $M_{ij}$  de même que les partitions des vecteurs  $\delta_U$ ,  $F$  et  $S$ .

D'autre part, on souhaite imposer une condition naturelle en  $x = 1$  de la forme :

$$q(1) \frac{du}{dx}(1) = d$$

ce qui revient à poser  $s_{12}^{K_3} = d$ . À l'autre extrémité, la condition naturelle est inconnue et donc  $s_{11}^{K_1}$  l'est aussi. De plus, puisqu'il n'y a aucune charge ponctuelle d'imposée, on a :

$$s_{12}^{K_1} + s_{11}^{K_2} = s_{12}^{K_2} + s_{11}^{K_3} = 0$$

On en déduit que :

$$S^I = [s_{11}^{K_1}]$$

et :

$$\delta_U^I = \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \end{bmatrix} \quad S^C = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ d \end{bmatrix}$$

Le relèvement des conditions aux limites est :

$$U^g = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ c \end{bmatrix}$$

On remarque immédiatement que le système est bien de la forme 5.19. La matrice  $M_{11}$  est de dimension 6 sur 6, la matrice  $M_{22}$  est de dimension 1 sur 1 et les matrices  $M_{12}$  et  $M_{21}$  de dimension 6 par 1 et 1 par 6 respectivement. La résolution est alors immédiate en se servant des relations 5.20 et 5.21, pourvu que l'on donne des valeurs précises à  $c$  et  $d$ .

Une fois le système résolu, la solution est  $u(x) = u_g(x) + \delta_u(x)$  qui sous forme vectorielle devient :

$$U = U^g + \delta_U = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ c \end{bmatrix} + \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ 0 \end{bmatrix} = \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ c \end{bmatrix}$$

◀

#### ◆ Remarque 5.7

Dans cet exemple, la matrice  $M_{11}$  possède une structure bien particulière. On remarque en effet que les coefficients non nuls sont concentrés le long des diagonales principales. Lorsque l'on s'éloigne des diagonales principales, les coefficients de la matrice sont nuls et on parle de *matrice bande* ou de *matrice en ligne de ciel*. Cela est dû à la numérotation initiale des degrés de liberté du domaine (voir la figure 5.3) qui met en relation (via le tableau d'adressage) un degré de liberté seulement avec les degrés de liberté adjacents. Une numérotation aléatoire des degrés de liberté aurait pour conséquence l'obtention d'une matrice pleine où les coefficients non nuls sont éparpillés dans toute la matrice. Il est important de profiter de cette structure particulière et de s'assurer que la numérotation des degrés de liberté minimise l'éparpillement des coefficients à l'intérieur de la matrice. Il existe plusieurs façons de faire et nous y reviendrons plus loin puisque cela prend toute son importance en dimension supérieure à 1. Notons enfin que cette renumérotation éventuelle n'a d'effets que sur le tableau *numer* (et par conséquent sur le tableau d'adressage *adres*) et que toute la procédure précédemment décrite reste inchangée. ◆

### 5.1.10 Présentation des résultats

La présentation des résultats est une opération importante mais qui n'a que peu de rapport avec la méthode des éléments finis en tant que telle. De nombreux logiciels commerciaux existent et permettent de présenter les résultats de façon claire et précise. La méthode des éléments finis et les méthodes numériques en général produisent des quantités impressionnantes de données que l'on souhaite interpréter. Si ce problème est relativement facile à surmonter en dimension 1, il en est tout autrement en dimension 2 ou 3.

Il convient toutefois de mentionner comment évaluer la solution calculée (et éventuellement ses dérivées si nécessaire) dans le but d'en faire un graphique.

- Pour chaque élément  $K$  ;
  - Lire la ligne  $K$  du tableau d'adressage ;
  - En déduire le vecteur  $U^K$  des inconnues élémentaires  $u_i^K$  ;
  - Pour chaque point  $x$  où on veut évaluer  $u^K$  ;
    - Sur l'élément de référence :  $\xi = \frac{2x - (x_1^K + x_2^K)}{h^K}$
    - $u^K(x) = \sum_{j=1}^{n_d^K} u_j^K \hat{\psi}_j(\xi)$  et  $\frac{du^K}{dx}(x) = \frac{2}{h^K} \sum_{j=1}^{n_d^K} u_j^K \frac{d\hat{\psi}_j}{d\xi}(\xi)$
- Fin de la boucle sur les éléments

### 5.1.11 Exemples et applications

Il est temps de s'attaquer à un exemple complet. Nous reviendrons donc sur les différentes étapes introduites jusqu'à maintenant et nous effectuerons une résolution complète.

#### ► Exemple 5.6

On considère la déflexion d'un câble de longueur 5m en tension entre 2 points d'attache tel qu'illustré à la figure 5.11. On exerce ainsi une tension de 400N sur le câble. Le poids linéaire du câble est de 6 N/m et de plus, on répartit une charge non uniforme de  $40x$  N/m dans l'intervalle  $[0, 2]$ . Enfin, en  $x = 4$ , une charge ponctuelle de 150 N contribue également à la déformation du câble. Pour résoudre ce problème, on doit considérer l'équation différentielle :

$$\begin{cases} -\frac{d}{dx} \left( T(x) \frac{du}{dx} \right) = f(x) \\ u(0) = 0, \quad u(5) = 0 \end{cases} \quad (5.22)$$

où  $T(x)$  est la tension dans le câble,  $u(x)$  est la déflexion verticale et  $f(x)$  est la charge appliquée. Dans ce cas précis, on applique une charge linéaire sur l'intervalle  $[0, 2]$ , en plus du poids du câble ce qui donne :

$$f(x) = \begin{cases} -(6 + 40x) & \text{pour } 0 \leq x \leq 2 \\ -6 & \text{pour } 2 \leq x \leq 5 \end{cases}$$

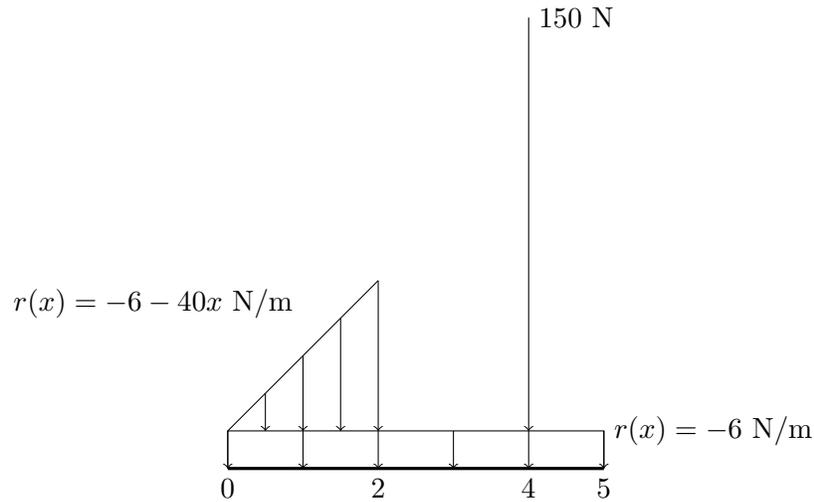


Figure 5.11 – Déflexion verticale d'un câble

Comme nous l'avons déjà souligné, la charge ponctuelle s'interprète comme le saut de la condition naturelle (de la variable secondaire) au point  $x = 4$  et sera prise en compte lors de l'imposition des conditions aux limites. Ce problème entre dans le cadre du problème type 5.1 en posant simplement  $p(x) = 0$ ,  $q(x) = T(x) = 400$  et  $r(x) = f(x)$ . Notons de plus que puisque les conditions aux limites essentielles sont homogènes, on peut prendre  $u_g(x) = 0$ .

– **Le maillage**

On considère un maillage de 4 éléments et une interpolation quadratique sur chaque élément. Le maillage est présenté aussi à la figure 5.12 ainsi que les coordonnées et la numérotation des noeuds. De plus, on s'est assuré de placer le point  $x = 4$ , où se situe la charge ponctuelle, à la frontière entre 2 éléments ce qui en facilite la prise en compte. On a donc les tableaux

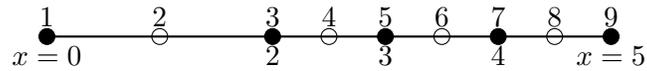


Figure 5.12 – Maillage pour la déflexion verticale d'un câble

suivants :

Coordonnées des noeuds		
Tableau <i>coor</i>		
$x_1$	$x_2$	$x_3$
0,0000	0,0000	0,0000
1,0000	0,0000	0,0000
2,0000	0,0000	0,0000
2,5000	0,0000	0,0000
3,0000	0,0000	0,0000
3,5000	0,0000	0,0000
4,0000	0,0000	0,0000
4,5000	0,0000	0,0000
5,0000	0,0000	0,0000

Numéros des noeuds de calcul			
Tableau <i>connec</i>			
Élément	Noeud #1	Noeud #2	Noeud #3
1	1	3	2
2	3	5	4
3	5	7	6
4	7	9	8

La variable essentielle  $u(x)$  est imposée en  $x = 0$  et en  $x = 5$ . On obtient toujours le tableau de numérotation en 2 étapes. Dans un premier temps, on a évité de numéroter les degrés de liberté associés aux noeuds 1 et 9 où des conditions essentielles sont imposées. Dans un deuxième temps, on numérote les degrés de liberté fixés :

Numérotation Tableau <i>numer</i>	
Noeud	<i>u</i>
1	?
2	1
3	2
4	3
5	4
6	5
7	6
8	7
9	?

→

Numérotation Tableau <i>numer</i>	
Noeud	<i>u</i>
1	8
2	1
3	2
4	3
5	4
6	5
7	6
8	7
9	9

Le tableau d'adressage correspondante est alors :

Numéros des dds Tableau <i>adres</i>			
Élément	Ddl #1	Ddl #2	Ddl #3
1	8	2	1
2	2	4	3
3	4	6	5
4	6	9	7

– **Le système élémentaire**

Le système élémentaire a pour coefficients :

$$a_{ij}^K = \frac{2}{h^K} \int_{-1}^1 400 \frac{d\hat{\psi}_j}{d\xi} \frac{d\hat{\psi}_i}{d\xi} d\xi$$

$$f_i^K = \frac{h^K}{2} \int_{-1}^1 f \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_i(\xi) d\xi$$

$$s_i^K = s_{11}^K \hat{\psi}_i(-1) + s_{12}^K \hat{\psi}_i(1)$$

Pour la matrice élémentaire, puisque dans ce cas on a  $p(x) = 0$  et  $q(x) = 400$ , on peut tirer profit de l'équation 5.16 pour obtenir :

$$\frac{400}{3h^K} \begin{bmatrix} 7 & 1 & -8 \\ 1 & 7 & -8 \\ -8 & -8 & 16 \end{bmatrix}$$

et ce quel que soit l'élément. Par contre, pour le calcul du terme de droite on distingue 2 cas suivant que l'on se place dans le premier élément ou dans les autres.

### 1. Premier élément

Dans le premier élément  $x_1^{K_1} = 0$ ,  $x_2^{K_1} = 2$  et donc  $h^{K_1} = 2$  et on doit évaluer :

$$\begin{aligned} f_i^{K_1} &= \frac{h^{K_1}}{2} \int_{-1}^1 f \left( \frac{(x_1^{K_1} + x_2^{K_1}) + h^{K_1} \xi}{2} \right) \hat{\psi}_i(\xi) d\xi \\ &= \frac{2}{2} \int_{-1}^1 f(1 + \xi) \hat{\psi}_i(\xi) d\xi \\ &= \int_{-1}^1 -(6 + 40(1 + \xi)) \hat{\psi}_i(\xi) d\xi \end{aligned}$$

On montre alors que :

$$f^{K_1} = \frac{1}{3} \begin{bmatrix} -6 \\ -86 \\ -184 \end{bmatrix}$$

Le vecteur d'adressage de cet élément est [8 2 1] et on trouve ainsi le système élémentaire :

$$\frac{200}{3} \left[ \begin{array}{c|ccc} & 8 & 2 & 1 \\ \hline 8 & 7 & 1 & -8 \\ 2 & 1 & 7 & -8 \\ 1 & -8 & -8 & 16 \end{array} \right] \begin{bmatrix} \delta_{u_1}^{K_1} \\ \delta_{u_2}^{K_1} \\ \delta_{u_3}^{K_1} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -6 \\ -86 \\ -184 \end{bmatrix} + \begin{bmatrix} s_{11}^{K_1} \\ s_{12}^{K_1} \\ 0 \end{bmatrix}$$

### 2. Les 3 autres éléments

Dans cet exemple précis, les 3 autres éléments sont identiques et vérifient  $h^K = 1$ . De plus, puisque  $r(x) = -6$ , on peut encore utiliser les relations 5.16 et 5.17 pour obtenir directement les systèmes élémentaires :

$$\frac{400}{3} \begin{bmatrix} 7 & 1 & -8 \\ 1 & 7 & -8 \\ -8 & -8 & 16 \end{bmatrix} \begin{bmatrix} \delta_{u_1}^K \\ \delta_{u_2}^K \\ \delta_{u_3}^K \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -3 \\ -3 \\ -12 \end{bmatrix} + \begin{bmatrix} s_{11}^K \\ s_{12}^K \\ 0 \end{bmatrix}$$

Seul le vecteur d'adressage varie pour ces 3 éléments.

## – L'assemblage

En se servant du tableau *adres*, on assemble (en exercice) la matrice 9 sur 9 suivante :

$$\frac{100}{3} \begin{bmatrix} 32 & -16 & 0 & 0 & 0 & 0 & 0 & -16 & 0 \\ -16 & 42 & -32 & 4 & 0 & 0 & 0 & 2 & 0 \\ 0 & -32 & 64 & -32 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & -32 & 56 & -32 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & -32 & 64 & -32 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & -32 & 56 & -32 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & -32 & 64 & 0 & -32 \\ -16 & 2 & 0 & 0 & 0 & 0 & 0 & 14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & -32 & 0 & 28 \end{bmatrix}$$

et le terme de droite :

$$\frac{1}{3} \begin{bmatrix} -184 \\ -89 \\ -12 \\ -6 \\ -12 \\ -6 \\ -12 \\ -6 \\ -3 \end{bmatrix} + \begin{bmatrix} 0 \\ s_{12}^{K_1} + s_{11}^{K_2} \\ 0 \\ s_{12}^{K_2} + s_{11}^{K_3} \\ 0 \\ s_{12}^{K_3} + s_{11}^{K_4} \\ 0 \\ s_{11}^{K_1} \\ s_{12}^{K_4} \end{bmatrix}$$

– **Conditions aux limites**

1. Conditions essentielles :

Deux degrés de liberté sont fixés par les conditions essentielles soient les degrés de liberté 8 et 9 pour lesquels  $\delta_{u_8} = \delta_{u_9} = 0$ .

2. Conditions naturelles :

Au point  $x = 4$ , on impose une charge ponctuelle ce qui se traduit par la condition naturelle :

$$s_{12}^{K_3} + s_{11}^{K_4} = -150$$

Tous les autres sauts de la variable secondaire sont nuls. De plus,  $s_{12}^{K_4}$  et  $s_{11}^{K_1}$  sont inconnus. On a maintenant :

$$S^C = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -150 \\ 0 \end{bmatrix} \text{ et } S^I = \begin{bmatrix} s_{11}^{K_1} \\ s_{12}^{K_4} \end{bmatrix}$$

– **Solution du système linéaire**

On doit donc résoudre le système global :

$$\frac{100}{3} \begin{bmatrix} 32 & -16 & 0 & 0 & 0 & 0 & 0 & -16 & 0 \\ -16 & 42 & -32 & 4 & 0 & 0 & 0 & 2 & 0 \\ 0 & -32 & 64 & -32 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & -32 & 56 & -32 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & -32 & 64 & -32 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & -32 & 56 & -32 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & -32 & 64 & 0 & -32 \\ \hline -16 & 2 & 0 & 0 & 0 & 0 & 0 & 14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & -32 & 0 & 28 \end{bmatrix} \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ \delta_{u_7} \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -184 \\ -89 \\ -12 \\ -6 \\ -12 \\ -6 \\ -12 \\ -6 \\ -3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -150 \\ 0 \\ s_{11}^{K_1} \\ s_{12}^{K_4} \end{bmatrix}$$

On effectue la résolution de l'équation 5.20 par une méthode de décomposition  $LU$ . La solution du système linéaire réduit de dimension 7 sur 7 donne le vecteur :

$$\delta_U^I = - \begin{bmatrix} 0,235\,000\,000 \\ 0,355\,000\,000 \\ 0,367\,708\,333 \\ 0,376\,666\,667 \\ 0,381\,875\,000 \\ 0,383\,333\,333 \\ 0,193\,541\,666 \end{bmatrix}$$

Le vecteur solution complet (en ajoutant les conditions essentielles) est donc :

$$U = U^g + \delta_U = - \begin{bmatrix} 0,235\,000\,000 \\ 0,355\,000\,000 \\ 0,367\,708\,333 \\ 0,376\,666\,667 \\ 0,381\,875\,000 \\ 0,383\,333\,333 \\ 0,193\,541\,666 \\ 0 \\ 0 \end{bmatrix}$$

La solution élément par élément est illustrée à la figure 5.13. On remarque que la solution  $u(x)$  est constituée de 4 polynômes de degré 2 alors que  $Tu'(x)$  est constituée de 4 polynômes linéaires (puisque  $T$  est constant). On remarque enfin que le saut des variables secondaires en  $x = 2$ ,  $x = 3$  n'est pas nul alors que celui en  $x = 4$  est près de 150. Cela démontre que les conditions aux limites naturelles ne sont imposées que faiblement. Un maillage plus fin permettrait d'obtenir éventuellement les bonnes valeurs de ces sauts.

On peut enfin récupérer les variables naturelles qui restent en se servant de la relation 5.21. On obtient alors les réactions :

$$S^I = \begin{bmatrix} s_{11}^{K_1} \\ s_{12}^{K_4} \end{bmatrix} = \frac{-100}{3} \begin{bmatrix} -16 & 2 & \cdots & 0 \\ 0 & 0 & \cdots & -32 \end{bmatrix} \begin{bmatrix} 0,235\,000\,000 \\ 0,355\,000\,000 \\ 0,367\,708\,333 \\ 0,376\,666\,667 \\ 0,381\,875\,000 \\ 0,383\,333\,333 \\ 0,193\,541\,666 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} -6 \\ -3 \end{bmatrix} = \begin{bmatrix} 103,67 \\ 156,33 \end{bmatrix}$$

exprimées en N et qui représentent les forces exercées aux 2 extrémités pour retenir le câble. La somme de ces 2 forces est de 260 N ce qui est exactement la charge totale exercée sur le câble :

$$6 \times 5 + 150 + \int_0^2 40x \, dx = 260N$$

◀

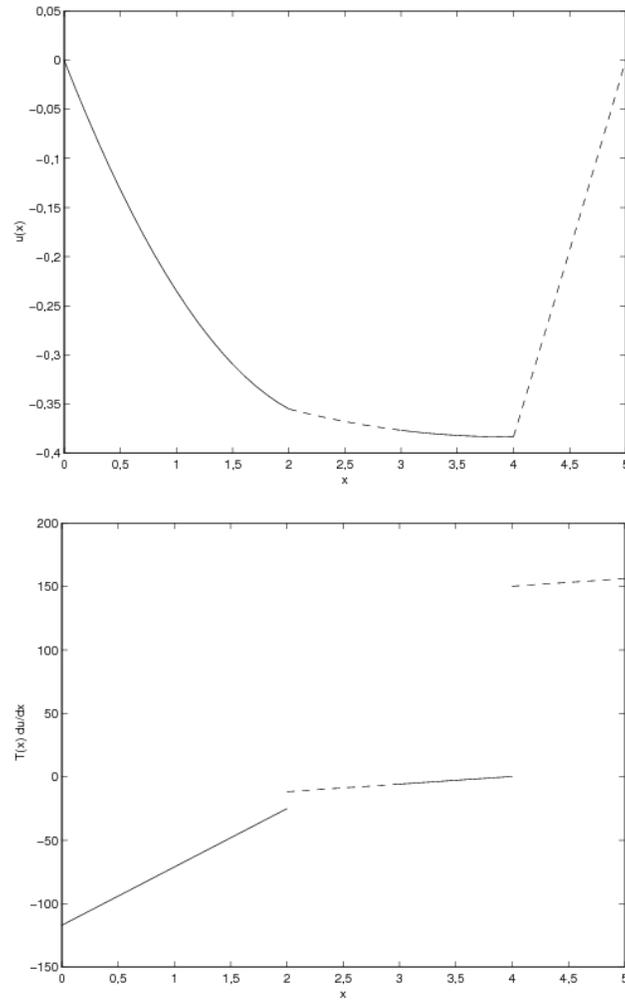
## 5.2 Équations différentielles d'ordre 4

Dans cette section nous nous intéressons aux équations différentielles d'ordre 4. Fort heureusement beaucoup de notions introduites à la section précédente restent inchangées et seront réutilisées avec très peu de modifications. C'est le cas du passage à l'élément de référence, de l'évaluation du système élémentaire, de l'assemblage, de l'imposition des conditions aux limites, etc. Seules quelques variantes propres aux équations différentielles d'ordre 4 sont nécessaires.

### 5.2.1 Problème type

Pour fixer les idées, nous commencerons par la résolution d'un problème classique d'une équation différentielle d'ordre 4 de la forme :

$$\left\{ \begin{array}{l} p(x)u + \frac{d^2}{dx^2} \left( q(x) \frac{d^2u}{dx^2} \right) = r(x) \\ u(0) = a, \quad q(0) \frac{d^2u}{dx^2}(0) = b \\ \frac{du}{dx}(L) = c, \quad \frac{d}{dx} \left( q(L) \frac{du}{dx}(L) \right) = d \end{array} \right. \quad (5.23)$$

Figure 5.13 – Solution numérique :  $u(x)$  et  $T(x)u'(x)$

Ce type d'équations différentielles est fréquemment rencontré pour les problèmes de poutres où dans ce cas,  $u(x)$  est la déflexion verticale de la poutre au point  $x$ . Nous savons d'ores et déjà que ce type de problèmes trouve un cadre fonctionnel approprié dans  $H^2(]0, L[)$  et qu'on distingue pour cette équation différentielle d'ordre 4 deux conditions essentielles (imposition de  $u(x)$  et de  $\frac{du}{dx}$ ) et deux conditions naturelles (imposition de  $\frac{d}{dx} \left( q(x) \frac{d^2u}{dx^2} \right)$  et de  $q(x) \frac{d^2u}{dx^2}$ ) en l'une ou l'autre des 2 extrémités du domaine. Tenant compte des conditions essentielles imposées dans ce problème, on pose :

$$V = \{w \in H^2(]0, L[) \mid w(0) = \frac{dw}{dx}(L) = 0\}$$

L'espace  $V$  est constitué des fonctions de  $H^2(]0, L[)$  qui s'annulent là où les conditions essentielles sont imposées. Dans ce cas, on pose  $u = \delta_u + u_g$  et la formulation variationnelle (laissée en exercice) est :

♠ trouver  $\delta_u \in V$  telle que :

$$a(\delta_u, w) = l(w) - a(u_g, w) \quad \forall w \in V \quad (5.24)$$

où :

$$a(\delta_u, w) = \int_0^1 (p(x)\delta_u(x)w(x) + q(x)\delta_u''(x)w''(x)) dx$$

et :

$$l(w) = \int_0^1 r(x)w(x)dx - d w(L) + b \frac{dw}{dx}(0)$$

La méthode de Ritz nécessite la construction d'un système linéaire global de la forme :

$$A\delta_U = F$$

où :

$$a_{ij} = a(\phi_j(x), \phi_i(x)) \quad \text{et} \quad f_i = l(\phi_i(x)) - a(u_g(x), \phi_i(x))$$

La fonction  $u_g(x)$  désigne comme toujours le relèvement des conditions aux limites essentielles. Il reste donc à construire les fonctions de Ritz  $\phi_i(x)$  dans  $H^2(\Omega)$  et qui seront polynômiales par élément.

### 5.2.2 Le maillage

Le maillage possède les mêmes caractéristiques que pour les problèmes d'ordre 2. Pour illustrer notre propos, on choisit comme domaine  $\Omega$  le segment  $[0, L]$  et on considère un maillage uniforme de 3 éléments. On prend seulement les 2 noeuds géométriques comme noeuds de calcul ( $n_g^K = n_c^K = 2$ ). Le tableau de connectivité des noeuds s'écrit dans ce cas :

Numéros des noeuds de calcul		
Tableau <i>conec</i>		
Élément	Noeud #1	Noeud #2
1	1	2
2	2	3
3	3	4

Rappelons encore ici qu'il y a 2 conditions essentielles pour ce problème et donc qu'il y a 2 degrés de liberté associés à chaque noeud d'interpolation ( $n_d^K = 4$ ). Nous continuerons à numéroter les degrés de liberté où des conditions essentielles sont imposées en tout dernier. Dans le cas qui nous intéresse, on impose  $u$  au premier noeud et on impose  $\frac{du}{dx}$  au quatrième noeud. On aurait la situation suivante :

Numérotation		
Tableau <i>numer</i>		
Noeud	Ddl #1 ( $u(x)$ )	Ddl #2 ( $\frac{du}{dx}$ )
1	?	1
2	2	3
3	4	5
4	6	?

Remarquons que le degré de liberté associé à la variable essentielle  $u(x)$  au premier noeud n'a pas été numéroté de même que celui associé à la variable essentielle  $\frac{du}{dx}$  au noeud # 4 ce qui est fait dans une deuxième étape :

Numérotation		
Tableau <i>numer</i>		
Noeud	Ddl #1 ( $u(x)$ )	Ddl #2 ( $\frac{du}{dx}$ )
1	7	1
2	2	3
3	4	5
4	6	8

On peut maintenant construire le tableau d'adressage :

Numéros des dds				
Tableau <i>adres</i>				
Élément	Ddl #1	Ddl #2	Ddl #3	Ddl #4
1	7	1	2	3
2	2	3	4	5
3	4	5	6	8

### 5.2.3 Formulation variationnelle élémentaire

Cette étape consiste à obtenir une formulation variationnelle sur un élément quelconque  $K$ . On intègre sur un élément  $K = [x_1^K, x_2^K]$  plutôt que sur le domaine  $\Omega$  au complet. On a alors après 2 intégrations par parties :

$$\begin{aligned} & \int_K \left( p(x)u(x)w(x) + q(x) \frac{d^2u}{dx^2} \frac{d^2w}{dx^2} \right) dx \\ &= \int_K r(x)w(x)dx + q(x) \frac{d^2u}{dx^2} \frac{dw}{dx} \Big|_{x_1^K}^{x_2^K} - \frac{d}{dx} \left( q(x) \frac{d^2u}{dx^2} \right) w(x) \Big|_{x_1^K}^{x_2^K} \end{aligned}$$

ou encore :

$$\begin{aligned} & \int_{x_1^K}^{x_2^K} \left( p(x)u(x)w(x) + q(x) \frac{d^2u}{dx^2} \frac{d^2w}{dx^2} \right) dx \\ &= \int_{x_1^K}^{x_2^K} r(x)w(x)dx + s_{12}^K w(x_2^K) + s_{11}^K w(x_1^K) + s_{22}^K \frac{dw}{dx}(x_2^K) + s_{21}^K \frac{dw}{dx}(x_1^K) \end{aligned}$$

où on a introduit les *variables secondaires* :

$$\begin{aligned} s_{11}^K &= \frac{d}{dx} \left( q(x) \frac{d^2u}{dx^2} \right) \Big|_{x=x_1^K} & s_{12}^K &= - \frac{d}{dx} \left( q(x) \frac{d^2u}{dx^2} \right) \Big|_{x=x_2^K} \\ s_{21}^K &= - \left( q(x) \frac{d^2u}{dx^2} \right) \Big|_{x=x_1^K} & s_{22}^K &= \left( q(x) \frac{d^2u}{dx^2} \right) \Big|_{x=x_2^K} \end{aligned}$$

Ces 2 variables secondaires correspondent aux 2 conditions naturelles aux 2 extrémités de l'élément. La notation à 2 indices indique que  $s_{i1}^K$  est la valeur de la  $i$  ième variable secondaire à la première extrémité de l'élément tandis que  $s_{i2}^K$  est la valeur de la  $i$  ième variable secondaire à la deuxième extrémité de l'élément  $K$ . Les signes qui sont affectés aux variables secondaires sont *a priori* arbitraires de manière à ce que les signes soient tous positifs dans la formulation variationnelle élémentaire. Mentionnons cependant qu'en dimension 2 ou 3, le signe est déterminé par le vecteur normal extérieur à la frontière de l'élément.

On passe ensuite en formulation de type correction en remplaçant  $u(x)$  par  $u_g(x) + \delta_u(x)$  :

$$\begin{aligned} & \int_{x_1^K}^{x_2^K} \left( p(x)\delta_u(x)w(x) + q(x) \frac{d^2\delta_u}{dx^2} \frac{d^2w}{dx^2} \right) dx \\ &= \int_{x_1^K}^{x_2^K} r(x)w(x)dx - \int_{x_1^K}^{x_2^K} \left( p(x)u_g(x)w(x) + q(x) \frac{d^2u_g}{dx^2} \frac{d^2w}{dx^2} \right) dx \\ &+ s_{12}^K w(x_2^K) + s_{11}^K w(x_1^K) + s_{22}^K \frac{dw}{dx}(x_2^K) + s_{21}^K \frac{dw}{dx}(x_1^K) \end{aligned}$$

On pose ensuite :

$$\delta_u(x) = \sum_{j=1}^{n_d^K} \delta_{u_j}^K \psi_j^K(x) \quad \text{et} \quad u_g(x) = \sum_{j=1}^{n_d^K} u_{g_j}^K \psi_j^K(x)$$

et on choisit ensuite successivement  $w(x) = \psi_i^K(x)$ ,  $i = 1, 2, \dots, n_d^K$  pour obtenir le système élémentaire :

$$A^K \delta_U^K = F^K + S^K$$

où :

$$\begin{aligned} a_{ij}^K &= \int_{x_1^K}^{x_2^K} \left( p(x) \psi_j^K(x) \psi_i^K(x) + q(x) \frac{d^2 \psi_j^K}{dx^2} \frac{d^2 \psi_i^K}{dx^2} \right) dx \\ f_i^K &= \int_{x_1^K}^{x_2^K} r(x) \psi_i^K(x) dx - \int_{x_1^K}^{x_2^K} \left( p(x) \left( \sum_{j=1}^{n_d^K} u_{g_j}^K \psi_j^K(x) \right) \psi_i^K(x) + q(x) \left( \sum_{j=1}^{n_d^K} u_{g_j}^K \frac{d^2 \psi_j^K}{dx^2} \right) \frac{d^2 \psi_i^K}{dx^2} \right) dx \\ s_i^K &= s_{12}^K \psi_i^K(x_2^K) + s_{11}^K \psi_i^K(x_1^K) + s_{22}^K \frac{d\psi_i^K}{dx}(x_2^K) + s_{21}^K \frac{d\psi_i^K}{dx}(x_1^K) \end{aligned}$$

#### 5.2.4 Passage à l'élément de référence

Effectuons maintenant le changement de variables dans le système élémentaire. Rappelons de l'équation 5.9 que :

$$\frac{d\psi_i^K}{dx}(x) = \frac{2}{h^K} \frac{d\hat{\psi}_i}{d\xi}(\xi), \quad \frac{d^2 \psi_i^K}{dx^2}(x) = \frac{4}{(h^K)^2} \frac{d^2 \hat{\psi}_i}{d\xi^2}(\xi) \quad \text{et} \quad dx = \frac{h^K}{2} d\xi$$

On obtient alors sur l'élément de référence :

$$\begin{aligned}
a_{ij}^K &= \int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) \frac{h^K}{2} d\xi \\
&\quad + \int_{-1}^1 q \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \left( \frac{d^2 \hat{\psi}_j}{d\xi^2} \frac{4}{(h^K)^2} \right) \left( \frac{d^2 \hat{\psi}_i}{d\xi^2} \frac{4}{(h^K)^2} \right) \frac{h^K}{2} d\xi \\
&= \frac{h^K}{2} \int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) d\xi \\
&\quad + \frac{8}{(h^K)^3} \int_{-1}^1 q \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \frac{d^2 \hat{\psi}_j}{d\xi^2} \frac{d^2 \hat{\psi}_i}{d\xi^2} d\xi \\
f_i^K &= \frac{h^K}{2} \int_{-1}^1 r \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_i(\xi) d\xi - \frac{h^K}{2} \int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \left( \sum_{j=1}^{n_d^K} u_{g_j}^K \hat{\psi}_j(\xi) \right) \hat{\psi}_i(\xi) d\xi \\
&\quad - \frac{8}{(h^K)^3} \int_{-1}^1 q \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \left( \sum_{j=1}^{n_d^K} u_{g_j}^K \frac{d^2 \hat{\psi}_j}{d\xi^2} \right) \frac{d^2 \hat{\psi}_i}{d\xi^2} d\xi \\
s_i^K &= s_{11}^K \hat{\psi}_i(-1) + s_{12}^K \hat{\psi}_i(1) + s_{21}^K \frac{2}{h^K} \frac{d\hat{\psi}_i}{d\xi}(-1) + s_{22}^K \frac{2}{h^K} \frac{d\hat{\psi}_i}{d\xi}(1)
\end{aligned}$$

Ce système est de dimension  $n_d^K$  sur  $n_d^K$ . Pour l'évaluer, il reste à construire des fonctions d'interpolation sur l'élément de référence. Puisque la solution recherchée est dans un sous-espace de  $H^2(\Omega)$ , l'approximation polynômiale devra être continue et dérivable au moins une fois (voir le chapitre 2). Cela nous amène à rappeler les polynômes d'Hermite.

### 5.2.5 Construction des fonctions d'interpolation

Pour simplifier l'exposé, nous commencerons par le cas où il y a seulement 2 noeuds de calcul qui correspondent aux 2 noeuds géométriques ( $n_c^K = n_g^K = 2$ ). Puisqu'il y a 2 conditions essentielles par noeud de calcul, on aura 2 degrés de liberté par noeud de calcul et donc 4 fonctions d'interpolation à construire ( $n_d^K = 4$ ). Sur un élément, on a donc :

$$u^K(x) = \sum_{j=1}^4 u_j^K \psi_j^K(x)$$

La valeur des inconnues  $u_j^K$  est définie par :

$$\begin{aligned} u^K(x_1^K) &= \sum_{j=1}^4 u_j^K \psi_j^K(x_1^K) = \sum_{j=1}^4 u_j^K \hat{\psi}_j(-1) = u_1^K \\ \frac{du^K}{dx}(x_1^K) &= \sum_{j=1}^4 u_j^K \frac{d\psi_j^K}{dx}(x_1^K) = \sum_{j=1}^4 u_j^K \frac{2}{h^K} \frac{d\hat{\psi}_j}{d\xi}(-1) = u_2^K \\ u^K(x_2^K) &= \sum_{j=1}^4 u_j^K \psi_j^K(x_2^K) = \sum_{j=1}^4 u_j^K \hat{\psi}_j(+1) = u_3^K \\ \frac{du^K}{dx}(x_2^K) &= \sum_{j=1}^4 u_j^K \frac{d\psi_j^K}{dx}(x_2^K) = \sum_{j=1}^4 u_j^K \frac{2}{h^K} \frac{d\hat{\psi}_j}{d\xi}(+1) = u_4^K \end{aligned}$$

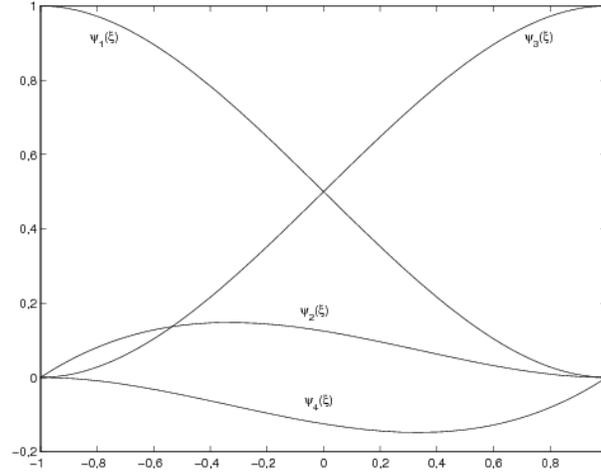
Les fonctions d'interpolation  $\hat{\psi}(\xi)$  doivent donc vérifier les conditions suivantes :

$$\begin{aligned} \hat{\psi}_1(-1) &= 1 \quad \text{et} \quad \hat{\psi}_i(-1) = 0 \quad \text{pour} \quad i \neq 1 \\ \frac{d\hat{\psi}_2}{d\xi}(-1) &= \frac{h^K}{2} \quad \text{et} \quad \frac{d\hat{\psi}_i}{d\xi}(-1) = 0 \quad \text{pour} \quad i \neq 2 \\ \hat{\psi}_3(+1) &= 1 \quad \text{et} \quad \hat{\psi}_i(+1) = 0 \quad \text{pour} \quad i \neq 3 \\ \frac{d\hat{\psi}_4}{d\xi}(+1) &= \frac{h^K}{2} \quad \text{et} \quad \frac{d\hat{\psi}_i}{d\xi}(+1) = 0 \quad \text{pour} \quad i \neq 4 \end{aligned}$$

Ces dernières propriétés ne sont rien de moins que la définition des polynômes d'Hermite (voir par exemple Burden et Faires, réf. [8]). Puisqu'il y a 4 conditions à satisfaire pour chacune des fonctions d'interpolation, il semble naturel d'utiliser des polynômes de degré 3 et il est alors facile de montrer que :

$$\begin{aligned} \hat{\psi}_1(\xi) &= \frac{1}{4}(1 - \xi)^2(2 + \xi) & \hat{\psi}_3(\xi) &= \frac{1}{4}(1 + \xi)^2(2 - \xi) \\ \hat{\psi}_2(\xi) &= \frac{h^K}{8}(1 - \xi^2)(1 - \xi) & \hat{\psi}_4(\xi) &= \frac{h^K}{8}(-1 + \xi^2)(1 + \xi) \end{aligned}$$

Ces fonctions sont illustrées à la figure 5.14 dans le cas particulier où  $h^K = 1$ .

Figure 5.14 – Fonctions d’interpolation d’Hermite ( $h^K = 1$ )

On montre ensuite facilement que :

$$\begin{aligned} \frac{d\hat{\psi}_1(\xi)}{d\xi} &= -\frac{3}{4}(1 - \xi^2) & \frac{d\hat{\psi}_3(\xi)}{d\xi} &= \frac{3}{4}(1 - \xi^2) \\ \frac{d\hat{\psi}_2(\xi)}{d\xi} &= \frac{h^K}{8}(-1 + \xi)(1 + 3\xi) & \frac{d\hat{\psi}_4(\xi)}{d\xi} &= \frac{h^K}{8}(-1 - \xi)(1 - 3\xi) \end{aligned}$$

tandis que les dérivées secondes s’écrivent :

$$\begin{aligned} \frac{d^2\hat{\psi}_1(\xi)}{d\xi^2} &= \frac{3\xi}{2} & \frac{d^2\hat{\psi}_3(\xi)}{d\xi^2} &= -\frac{3\xi}{2} \\ \frac{d^2\hat{\psi}_2(\xi)}{d\xi^2} &= \frac{h^K}{4}(-1 + 3\xi) & \frac{d^2\hat{\psi}_4(\xi)}{d\xi^2} &= \frac{h^K}{4}(1 + 3\xi) \end{aligned}$$

Notons que la définition des fonctions d’interpolation sur l’élément de référence dépend de la longueur  $h^K$  de l’élément  $K$ , ce qui n’était nullement le cas pour les fonctions d’interpolation de Lagrange. Ceci est dû à la présence de degrés de liberté portant sur la dérivée de  $u(x)$ . Pour assurer la continuité de la dérivée à la frontière entre 2 éléments, il est nécessaire d’ajuster les fonctions d’interpolation en fonction de la longueur des éléments adjacents à un noeud.

En utilisant ces fonctions d’interpolation, on donne une signification physique précise au vecteur des degrés de liberté élémentaire et par la suite, au vecteur global des degrés de liberté du domaine.

Ainsi, le vecteur  $U^K$  correspond aux valeurs de  $u(x)$  et de  $\frac{du}{dx}$ , les 2 variables essentielles aux noeuds de calcul c.-à-d.

$$\begin{aligned} u_1^K &= u(x_1^K) & u_3^K &= u(x_2^K) \\ u_2^K &= \frac{du}{dx}(x_1^K) & u_4^K &= \frac{du}{dx}(x_2^K) \end{aligned}$$

◆ **Remarque 5.8**

La notation utilisée peut peut-être porter à confusion. En effet, le degré de liberté  $u_2^K$  contient la valeur de  $\frac{du}{dx}(x_1^K)$  et non la valeur de  $u$  au noeud correspondant. Nous aurions pu utiliser une notation différente mais il suffit de se rappeler que  $u_2^K$  correspond au deuxième degré de liberté de l'élément  $K$  et que son interprétation peut varier suivant le problème. ◆

### 5.2.6 Évaluation du système élémentaire

L'évaluation du système élémentaire suit les mêmes lignes que pour les problèmes d'ordre 2. On peut toujours choisir entre l'intégration numérique et l'intégration exacte. Si on suppose encore une fois que les fonctions  $p(x)$ ,  $q(x)$  et  $r(x)$  sont constantes par élément (et  $u_g(x) = 0$ ), le système élémentaire devient :

$$A^K = \frac{h^K p^K}{13440} \begin{bmatrix} 4992 & 704h^K & 1728 & -416h^K \\ 704h^K & 128(h^K)^2 & 416h^K & -96(h^K)^2 \\ 1728 & 416h^K & 4992 & -704h^K \\ -416h^K & -96(h^K)^2 & -704h^K & 128(h^K)^2 \end{bmatrix} + \frac{2q^K}{(h^K)^3} \begin{bmatrix} 6 & 3h^K & -6 & 3h^K \\ 3h^K & 2(h^K)^2 & -3h^K & (h^K)^2 \\ -6 & -3h^K & 6 & -3h^K \\ 3h^K & (h^K)^2 & -3h^K & 2(h^K)^2 \end{bmatrix} \quad (5.25)$$

Une intégration par une quadrature de Gauss-Legendre nécessiterait au moins 4 points d'intégration car la première matrice de l'équation 5.25 requiert l'intégration de polynômes de degré 6. En ce qui concerne le membre de droite, l'intégrand est cubique et une formule de Gauss-Legendre à 2 points suffirait. On a :

$$f^K + s^K = \frac{h^K r^K}{12} \begin{bmatrix} 6 \\ h^K \\ 6 \\ -h^K \end{bmatrix} + \begin{bmatrix} s_{11}^K \\ s_{21}^K \\ s_{12}^K \\ s_{22}^K \end{bmatrix} \quad (5.26)$$

### 5.2.7 Assemblage

Rien de neuf sous le soleil en ce qui concerne l'assemblage. On utilise encore le tableau d'adressage *adres* préalablement construit et contenant pour chaque élément les numéros de ses degrés de liberté. On procède exactement comme à la section 5.1.7 pour assembler.

### 5.2.8 Imposition des conditions aux limites

L'imposition des conditions aux limites se fait toujours en 2 étapes. Premièrement les degrés de liberté imposés par les conditions essentielles sont regroupés encore ici dans le vecteur  $U^C$ . On devra donc calculer le vecteur  $U^I$ .

Puisqu'il y a 2 conditions naturelles, le vecteur  $S$  fait intervenir deux types de sauts à la frontière de 2 éléments  $K^+$  et  $K^-$ . Si on note  $x$  le noeud de calcul correspondant à cette frontière et  $i$  et  $j$  les numéros des degrés de liberté associés à ce noeud, le vecteur  $S$  contiendra les coefficients :

$$s_i = s_{12}^{K^-} + s_{11}^{K^+} \quad \text{et} \quad s_j = s_{22}^{K^-} + s_{21}^{K^+}$$

Les indices  $-$  et  $+$  réfèrent aux valeurs à gauche et à droite de la variable en  $x$ . Ces expressions désignent les sauts des 2 variables secondaires au noeud correspondant.

Or si l'un ou l'autre de ces sauts était différent de 0, l'équation différentielle ferait apparaître une distribution de Dirac  $H_1\delta_x$  ou bien une dérivée d'une distribution de Dirac  $H_2\delta'_x$ . Si tel est le cas, c'est au moment de l'imposition des conditions naturelles à ce noeud que l'on prend en compte la contribution de cette charge ponctuelle et on pose :

$$s_i = s_{12}^{K^-} + s_{11}^{K^+} = H_1 \quad \text{et/ou} \quad s_j = s_{22}^{K^-} + s_{21}^{K^+} = H_2$$

Sinon, on pose simplement  $s_i = s_j = 0$ .

### 5.2.9 Solution du système global

On procède comme à la section 5.1.9 pour résoudre le système.

### 5.2.10 Exemples et applications

Nous allons considérer une poutre de longueur  $L$  m encastée à l'une de ses extrémités tel qu'illustrée à la figure 5.15. Ce problème cadre parfaitement avec notre problème type. Les variables essentielles  $u(x)$  et  $u'(x)$  désignent dans ce cas le déplacement vertical (positif si dirigé vers le haut) et la pente. Il suffit ensuite de poser  $p(x) = 0$  et  $q(x) = EI$ , le produit du module de Young  $E$  (en  $\text{N/m}^2$ ) et du moment d'inertie  $I$  de la poutre (en  $\text{m}^4$ ). Enfin,  $r(x)$  est une charge répartie sur la poutre (la charge est positive si dirigée vers le haut). Les 2 conditions naturelles désignent respectivement un effort tranchant (cisaillement) :

$$F(x) = \frac{d}{dx} \left( q(x) \frac{d^2u}{dx^2} \right) = \frac{d}{dx} \left( EI \frac{d^2u}{dx^2} \right) \quad \text{en N}$$

et un moment (couple) de flexion :

$$M(x) = \left( q(x) \frac{d^2u}{dx^2} \right) = \left( EI \frac{d^2u}{dx^2} \right) \quad \text{en N}\cdot\text{m}$$

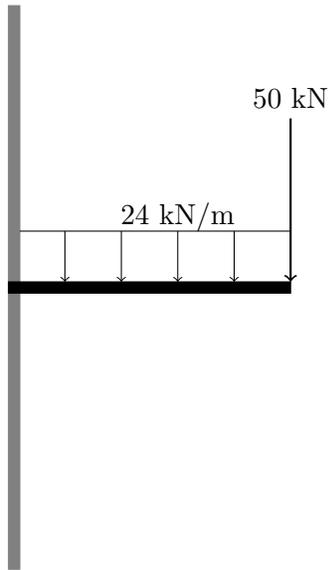


Figure 5.15 – Problème de la poutre encastrée

Rappelons que l'équation différentielle s'écrit alors :

$$\frac{d}{dx}F(x) = \frac{d^2}{dx^2}M(x) = r(x) \quad (5.27)$$

que l'on peut interpréter au sens des distributions. Dans un premier temps, si  $M(x)$  possède une discontinuité de première espèce au point  $x$ , alors la relation 5.27 nous assure que l'on verra apparaître la dérivée de la distribution de Dirac  $\delta'_x$  puisqu'on doit dériver 2 fois cette discontinuité. Un couple de flexion en  $x$  est donc modélisé par la distribution  $\delta'_x$ . Par contre, si la force de cisaillement  $F(x)$  possède un saut au point  $x$ , alors seulement la distribution de Dirac  $\delta_x$  apparaîtra.

Nous supposons dans cet exemple que le produit  $EI$  est constant et qu'une charge uniforme  $r(x) = r_0$  N/m est répartie sur toute la poutre. Enfin, à la deuxième extrémité ( $x = L$ ), on place une charge ponctuelle de  $f_0$  N. On remarque qu'on impose les deux variables essentielles  $u(0) = 0$  et  $\frac{du}{dx}(0) = 0$  à la première extrémité traduisant ainsi le fait que la poutre est encastrée à cet endroit. On choisira donc  $u_g(x) = 0$  et nous suivrons encore ici les étapes de résolution déjà décrites.

– **Le maillage**

On prendra pour ce problème un maillage très simple de 2 éléments de longueur égale  $h = L/2$  m et comportant 3 noeuds (*nnoeuds* = 3). Dans un premier temps, on obtient le tableau

*coord* :

<b>Coordonnées des noeuds</b>		
Tableau <i>coord</i>		
$x_1$	$x_2$	$x_3$
0	0	0
$L/2$	0	0
$L$	0	0

de même que le tableau de connectivité :

<b>Numéros des noeuds de calcul</b>		
Tableau <i>connec</i>		
Élément	Noeud #1	Noeud #2
1	1	2
2	2	3

Par la suite, on peut numéroter les degrés de liberté associés à chaque noeud. Pour les problèmes d'ordre 4, il y a 2 degrés de liberté par noeud et on a dans un premier temps :

<b>Numérotation</b>		
Tableau <i>numer</i>		
Noeud	Ddl #1 ( $u(x)$ )	Ddl #2 ( $\frac{du}{dx}$ )
1	?	?
2	1	2
3	3	4

puisque les 2 degrés de liberté associés au premier noeud sont fixés par les conditions essentielles. Un deuxième passage nous donne maintenant :

<b>Numérotation</b>		
Tableau <i>numer</i>		
Noeud	Ddl #1 ( $u(x)$ )	Ddl #2 ( $\frac{du}{dx}$ )
1	5	6
2	1	2
3	3	4

On peut maintenant construire le tableau d'adressage :

Numéros des dds				
Tableau <i>adres</i>				
Élément	Ddl #1	Ddl #2	Ddl #3	Ddl #4
1	5	6	1	2
2	1	2	3	4

Rappelons que pour un élément  $K$  quelconque :

$$\begin{aligned} \text{adres}(K, 1) &= \text{numer}(\text{conec}(K, 1), 1) \\ \text{adres}(K, 2) &= \text{numer}(\text{conec}(K, 1), 2) \\ \text{adres}(K, 3) &= \text{numer}(\text{conec}(K, 2), 1) \\ \text{adres}(K, 4) &= \text{numer}(\text{conec}(K, 2), 2) \end{aligned}$$

et ainsi, si on choisit l'élément  $K = 1$  on trouve :

$$\begin{aligned} \text{adres}(1, 1) &= \text{numer}(\text{conec}(1, 1), 1) = 5 \\ \text{adres}(1, 2) &= \text{numer}(\text{conec}(1, 1), 2) = 6 \\ \text{adres}(1, 3) &= \text{numer}(\text{conec}(1, 2), 1) = 1 \\ \text{adres}(1, 4) &= \text{numer}(\text{conec}(1, 2), 2) = 2 \end{aligned}$$

#### – Le système élémentaire

Les propriétés étant constantes dans tout le domaine, on utilise avantageusement les équations 5.25 et 5.26 pour obtenir immédiatement le système élémentaire :

$$\frac{2q^K}{(h^K)^3} \begin{bmatrix} 6 & 3h^K & -6 & 3h^K \\ 3h^K & 2(h^K)^2 & -3h^K & (h^K)^2 \\ -6 & -3h^K & 6 & -3h^K \\ 3h^K & (h^K)^2 & -3h^K & 2(h^K)^2 \end{bmatrix} = \frac{h^K r^K}{12} \begin{bmatrix} 6 \\ h^K \\ 6 \\ -h^K \end{bmatrix} + \begin{bmatrix} s_{11}^K \\ s_{21}^K \\ s_{12}^K \\ s_{22}^K \end{bmatrix}$$

où  $h^K = L/2$ ,  $q^K = EI$ , et  $r^K = r_0$ .

#### – L'assemblage

Si on pose maintenant  $L = 2$  ( $h^K = 1$ ),  $EI = q^K = 5000$  kN m<sup>2</sup>,  $r_0 = r^K = -24$  kN/m et  $f_0 = -50$  kN, les systèmes élémentaires sur les 2 éléments sont les mêmes. Le tableau d'adressage nous permet d'écrire pour le premier élément :

$$10\,000 \begin{bmatrix} \hline 5 & 6 & 1 & 2 \\ \hline 5 & 6 & 3 & -6 & 3 \\ 6 & 3 & 2 & -3 & 1 \\ 1 & -6 & -3 & 6 & -3 \\ 2 & 3 & 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} \overline{\delta_{u_1}^{K_1}} \\ \overline{\delta_{u_2}^{K_1}} \\ \overline{\delta_{u_3}^{K_1}} \\ \overline{\delta_{u_4}^{K_1}} \end{bmatrix} = \begin{bmatrix} -12 \\ -2 \\ -12 \\ 2 \end{bmatrix} + \begin{bmatrix} \overline{s_{11}^{K_1}} \\ \overline{s_{21}^{K_1}} \\ \overline{s_{12}^{K_1}} \\ \overline{s_{22}^{K_1}} \end{bmatrix}$$

tandis que sur le deuxième élément, on a :

$$10\,000 \left[ \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 6 & 3 & -6 & 3 \\ 2 & 3 & 2 & -3 & 1 \\ 3 & -6 & -3 & 6 & -3 \\ 4 & 3 & 1 & -3 & 2 \end{array} \right] \begin{bmatrix} \delta_{u_1}^{K_2} \\ \delta_{u_2}^{K_2} \\ \delta_{u_3}^{K_2} \\ \delta_{u_4}^{K_2} \end{bmatrix} = \begin{bmatrix} -12 \\ -2 \\ -12 \\ 2 \end{bmatrix} + \begin{bmatrix} s_{11}^{K_2} \\ s_{21}^{K_2} \\ s_{12}^{K_2} \\ s_{22}^{K_2} \end{bmatrix}$$

Le vecteur d'adressage du premier élément  $K_1$  est  $[5, 6, 1, 2]$ . Cela signifie que les inconnues élémentaires  $u_i^{K_1}$  correspondent aux degrés de liberté  $u_5, u_6, u_1$  et  $u_2$  du système global qui est de dimension 6 sur 6 dans ce cas. Le système élémentaire est donc équivalent à :

$$10\,000 \begin{bmatrix} 6 & -3 & 0 & 0 & -6 & -3 \\ -3 & 2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -6 & 3 & 0 & 0 & 6 & 3 \\ -3 & 1 & 0 & 0 & 3 & 2 \end{bmatrix} \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \end{bmatrix} = \begin{bmatrix} -12 \\ 2 \\ 0 \\ 0 \\ -12 \\ -2 \end{bmatrix} + \begin{bmatrix} s_{12}^{K_1} \\ s_{22}^{K_1} \\ 0 \\ 0 \\ s_{11}^{K_1} \\ s_{21}^{K_1} \end{bmatrix}$$

tandis que sur le deuxième élément on a :

$$10\,000 \begin{bmatrix} 6 & 3 & -6 & 3 & 0 & 0 \\ 3 & 2 & -3 & 1 & 0 & 0 \\ -6 & -3 & 6 & -3 & 0 & 0 \\ 3 & 1 & -3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \end{bmatrix} = \begin{bmatrix} -12 \\ -2 \\ -12 \\ 2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} s_{11}^{K_2} \\ s_{21}^{K_2} \\ s_{12}^{K_2} \\ s_{22}^{K_2} \\ 0 \\ 0 \end{bmatrix}$$

Si on additionne les contributions, on obtient le système global :

$$10\,000 \left[ \begin{array}{ccc|cc} 12 & 0 & -6 & 3 & -6 & -3 \\ 0 & 4 & -3 & 1 & 3 & 1 \\ -6 & -3 & 6 & -3 & 0 & 0 \\ 3 & 1 & -3 & 2 & 0 & 0 \\ \hline -6 & 3 & 0 & 0 & 6 & 3 \\ -3 & 1 & 0 & 0 & 3 & 2 \end{array} \right] \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \end{bmatrix} = \begin{bmatrix} -24 \\ 0 \\ -12 \\ 2 \\ -12 \\ -2 \end{bmatrix} + \begin{bmatrix} s_{11}^{K_2} + s_{12}^{K_1} \\ s_{21}^{K_2} + s_{22}^{K_1} \\ s_{12}^{K_2} \\ s_{22}^{K_2} \\ s_{11}^{K_1} \\ s_{21}^{K_1} \end{bmatrix}$$

#### – Les conditions aux limites

Deux conditions essentielles sont imposées :  $\delta_{u_5} = \delta_{u_6} = 0$ . Par contre, un cisaillement est imposé au noeud 3. Cela signifie que :

$$s_{12}^{K_2} = -50$$

et puisqu'aucun couple de flexion n'est imposé, on a :

$$s_{22}^{K_2} = 0$$

De plus, puisqu'aucun cisaillement ni couple de flexion n'est imposé au noeud 2, on a :

$$s_{11}^{K_2} + s_{12}^{K_1} = 0 \text{ et } s_{21}^{K_2} + s_{22}^{K_1} = 0$$

– **Résolution du système linéaire**

Après simplification, il reste le système linéaire :

$$10\,000 \begin{bmatrix} 12 & 0 & -6 & 3 \\ 0 & 4 & -3 & 1 \\ -6 & -3 & 6 & -3 \\ 3 & 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} \delta u_1 \\ \delta u_2 \\ \delta u_3 \\ \delta u_4 \end{bmatrix} = \begin{bmatrix} -24 \\ 0 \\ -12 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -50 \\ 0 \end{bmatrix}$$

En résolvant ce système et en ajoutant les conditions essentielles ( $U^g = 0$ ) qui étaient imposées, on obtient la solution complète :

$$\begin{bmatrix} \delta u_1 \\ \delta u_2 \\ \delta u_3 \\ \delta u_4 \\ \delta u_5 \\ \delta u_6 \end{bmatrix} = \begin{bmatrix} -1,173\,333 \times 10^{-2} \\ -2,060\,000 \times 10^{-2} \\ -3,626\,667 \times 10^{-2} \\ -2,640\,000 \times 10^{-2} \\ 0 \\ 0 \end{bmatrix}$$

Puisque  $u_g(x) = 0$ , on a  $u(x) = \delta_u(x)$  et dans le premier élément, le tableau d'adressage nous permet d'écrire la solution :

$$\begin{aligned} u^{K_1}(x) &= \sum_{j=1}^4 u_j^{K_1} \psi_j^{K_1}(x) = u_5 \psi_1^{K_1}(x) + u_6 \psi_2^{K_1}(x) + u_1 \psi_3^{K_1}(x) + u_2 \psi_4^{K_1}(x) \\ &= -10^{-2} \left( 1,173\,333 \psi_3^{K_1}(x) + 2,060\,000 \psi_4^{K_1}(x) \right) \end{aligned}$$

De même sur le deuxième élément, on trouve :

$$\begin{aligned} u^{K_2}(x) &= \sum_{j=1}^4 u_j^{K_2} \psi_j^{K_2}(x) = u_1 \psi_1^{K_2}(x) + u_2 \psi_2^{K_2}(x) + u_3 \psi_3^{K_2}(x) + u_4 \psi_4^{K_2}(x) \\ &= -10^{-2} \left( 1,173\,333 \psi_1^{K_2}(x) + 2,060\,000 \psi_2^{K_2}(x) + 3,626\,667 \psi_3^{K_2}(x) + 2,640,000 \psi_4^{K_2}(x) \right) \end{aligned}$$

Il ne reste qu'à récupérer les variables secondaires à la première extrémité de la poutre qui rappelons-le, n'étaient pas imposées. On a :

$$\begin{bmatrix} s_{11}^{K_1} \\ s_{21}^{K_1} \end{bmatrix} = 10\,000 \begin{bmatrix} -6 & 3 & 0 & 0 \\ -3 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta u_1 \\ \delta u_2 \\ \delta u_3 \\ \delta u_4 \end{bmatrix} + \begin{bmatrix} 12 \\ 2 \end{bmatrix} = \begin{bmatrix} 98,0 \\ 148,0 \end{bmatrix}$$

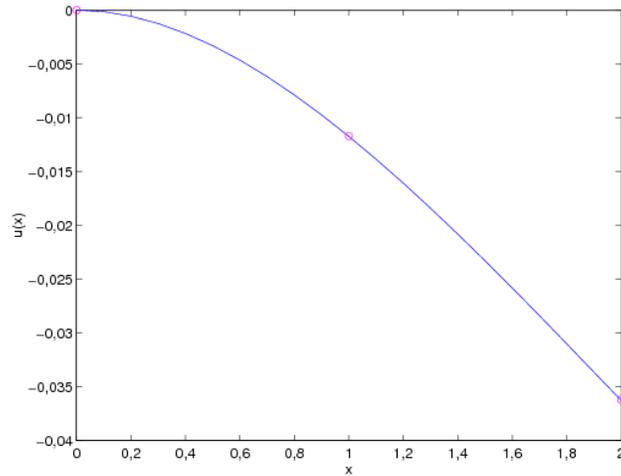


Figure 5.16 – Déformation de la poutre encastree

Ce problème possède une solution analytique dans le cas plus général où on impose également un couple de flexion  $M_0$  à l'extrémité droite de la poutre :

$$\begin{aligned}
 u(x) &= \left( \frac{1}{24}r_0x^4 - \frac{1}{6}(f_0 + r_0L)x^3 + \frac{1}{2}(M_0 + f_0L + \frac{1}{2}r_0L^2)x^2 \right) / EI \\
 u'(x) &= \left( \frac{1}{6}r_0x^3 - \frac{1}{2}(f_0 + r_0L)x^2 + (M_0 + f_0L + \frac{1}{2}r_0L^2)x \right) / EI \\
 EIU''(x) &= \left( \frac{1}{2}r_0x^2 - (f_0 + r_0L)x + (M_0 + f_0L + \frac{1}{2}r_0L^2) \right) \\
 EIU'''(x) &= r_0x - (f_0 + r_0L)
 \end{aligned}$$

Il suffit de prendre  $M_0 = 0$  dans le cas qui nous intéresse. On présente les solutions analytiques et numériques à la figure 5.16.

### 5.3 Exercices

1. Dans le cas quadratique, on peut utiliser une base dite hiérarchique pour interpoler. On pose alors en vertu de l'équation 5.4 :

$$u(x) |_K \simeq u^K(x) = \sum_{j=1}^{n_D} u_j^K \psi_j^K(x)$$

où sur l'élément de référence, on pose  $\hat{\psi}_1(\xi) = (1 - \xi)/2$ ,  $\hat{\psi}_2(\xi) = (1 + \xi)/2$  et  $\hat{\psi}_3(\xi) = (1 - \xi^2)$  c.-à-d. des fonctions linéaires sur les noeuds géométriques et une fonction quadratique au noeud milieu. Interpréter les  $u_j^K$  en fonction des valeurs aux noeuds de calcul  $u^K(x_i^K)$ .

2. Faire l'assemblage de la matrice et du système global de la page 116 en utilisant directement les systèmes élémentaires calculés.
3. Obtenir les fonctions d'interpolation d'Hermite de la page 125.
4. En vous servant des données de l'exemple 5.6 des notes de cours :
- a) Calculer :

$$-T(x) \frac{du^{K_1}}{dx} \Big|_{x=0} = -400 \frac{du^{K_1}}{dx}(0)$$

en utilisant le passage à l'élément de référence et comparer avec la valeur  $s_{11}^{K_1} = 103,67$  que nous avons obtenue. Commenter sur le fait que ces deux quantités sont des approximations de  $-T(x) \frac{du}{dx} \Big|_{x=0}$  dont la valeur exacte est 104.

- b) Si on remplace les conditions aux limites actuelles ( $u(0) = u(5) = 0$ ) de ce problème par  $u(0) = 2$  et  $u(5) = -3$ , donner l'expression de la fonction de relèvement des conditions aux limites ( $u_g$ ) qui sera implicitement construite par notre méthode d'éléments finis. Dessiner-cette fonction sur le maillage de l'exemple 5.6.
5. La formulation variationnelle d'une équation différentielle fait intervenir un terme de la forme :

$$\int_{x_1^K}^{x_2^K} c_2 \frac{du}{dx}(x) w(x) dx$$

où  $c_2$  est une constante. Si on utilise des éléments linéaires, la matrice élémentaire associée à ce terme est de dimension 2 par 2. Évaluer la matrice élémentaire. **On passera bien sûr par l'élément de référence.**

6. Soit le problème suivant :

$$\begin{cases} -\frac{d}{dx} \left( x^3 \frac{du}{dx} \right) + x^2 u & = \frac{1}{x+1} \\ u(0) & = -1 \\ u(1) & = 3 \end{cases}$$

On considérera un maillage uniforme de 3 éléments de l'intervalle  $]0, 1[$  et on utilisera une approximation cubique dans chaque élément.

- a) Déterminer la fonction d'interpolation  $\hat{\psi}_3$  dans l'élément de référence.
- b) Dessiner le relèvement des conditions aux limites essentielles sur le maillage utilisé ;
- c) Déterminer le nombre de points de Gauss minimum afin d'évaluer exactement chaque terme de la formulation variationnelle ;
- d) Donner la table d'adressage correspondant au maillage utilisé ;
- e) Déterminer, en fonction des matrices élémentaires  $a_{ij}^K$ , la valeur de  $a_{33}$  et de  $a_{14}$  du système global.

## Chapitre 6

# Éléments finis multidimensionnels

Notre présentation de la méthode des éléments finis unidimensionnels était suffisamment générale pour passer facilement au cas multidimensionnel. Bien sûr quelques difficultés supplémentaires se présenteront mais rien de franchement nouveau. Il faudra par exemple accorder une attention accrue à la numérotation des degrés de liberté pour éviter des matrices exigeant trop d'espace-mémoire. Il est de plus évident que la taille des problèmes augmentera de manière significative, surtout en dimension 3.

### 6.1 Problème type

Nous considérerons l'équation aux dérivées partielles suivante :

$$\left\{ \begin{array}{l} p(\mathbf{x})u(\mathbf{x}) - \nabla \cdot (q(\mathbf{x})\nabla u) = r(\mathbf{x}) \text{ dans } \Omega \\ u(\mathbf{x}) = g(\mathbf{x}) \text{ sur } \Gamma_0 \\ q(\mathbf{x})\frac{\partial u}{\partial \mathbf{n}} = h(\mathbf{x}) \text{ sur } \Gamma_1 \end{array} \right. \quad (6.1)$$

On suppose que la fonction  $p(\mathbf{x})$  est positive ou nulle et que  $0 < q_1 \leq q(\mathbf{x}) \leq q_2$  sur le domaine  $\Omega$ . Nous supposerons de plus que les frontières  $\Gamma_0$  et  $\Gamma_1$  constituent une partition de la frontière  $\Gamma$  du domaine c.-à-d.  $\Gamma_0 \cap \Gamma_1 = \emptyset$  et  $\Gamma_0 \cup \Gamma_1 = \Gamma$ . Il faut de plus spécifier la régularité des conditions aux limites. Puisqu'il s'agit d'un problème d'ordre 2, on travaillera dans  $H^1(\Omega)$  et l'imposition de  $u(\mathbf{x})$  sera l'unique condition essentielle. Il s'en suit que l'espace fonctionnel approprié est l'espace  $H_{\Gamma_0}^1(\Omega)$ . On doit donc supposer que  $g(\mathbf{x}) \in H^{1/2}(\Gamma_0)$  ce qui nous permettra d'effectuer le relèvement de la condition aux limites essentielle par une fonction  $u_g(\mathbf{x})$  de  $H^1(\Omega)$ .

Sous ces conditions, les hypothèses du théorème de Lax-Milgram sont vérifiées. On obtient de plus la formulation variationnelle (laissée en exercice) :

♠ Trouver  $\delta_u(\mathbf{x}) \in H_{\Gamma_0}^1(\Omega)$  telle que :

$$a(\delta_u, w) = l(w) - a(u_g, w) \quad \forall w(\mathbf{x}) \in H_{\Gamma_0}^1(\Omega)$$

où :

$$a(\delta_u, w) = \int_{\Omega} (p(\mathbf{x})\delta_u(\mathbf{x})w(\mathbf{x}) + q(\mathbf{x})\nabla\delta_u \cdot \nabla w) \, dv$$

et :

$$l(w) = \int_{\Omega} r(\mathbf{x})w(\mathbf{x}) \, dv + \int_{\Gamma_1} h(\mathbf{x})w(\mathbf{x}) \, ds$$

Ce problème généralise naturellement le problème 5.1 du chapitre 5.

## 6.2 Le maillage

Il est de toute évidence plus difficile de concevoir un maillage en dimension 2 ou 3. En effet, découper un domaine de forme quelconque en sous-domaines de formes géométriques simples n'est pas une mince tâche. Plusieurs techniques de génération de maillage existent avec des avantages et inconvénients de toutes sortes. Nous ne mentionnerons à ce sujet que le mailleur GMSH disponible gratuitement à l'adresse :

<http://geuz.org/gmsh/>

qui produit des maillages de triangles en dimension 2 et de tétraèdres en dimension 3.

Dans la majeure partie de ce chapitre, nous nous limiterons à des maillages très simples comportant peu d'éléments. Les applications de fin de chapitre permettront au lecteur de voir des exemples concrets et relativement réalistes.

Avant de parler de maillage, il faut décider de la forme géométrique des éléments. En dimension 1, ce choix est facile mais déjà en dimension 2, des variantes existent. Il faut choisir une forme géométrique simple. En dimension 2 nous considérerons les triangles et les quadrilatères tandis qu'en dimension 3 nous nous limiterons aux tétraèdres. En fait, la vaste majorité des mailleurs modernes produisent des triangles en dimension 2 et des tétraèdres en dimension 3. Il est en effet beaucoup plus difficile de décomposer un domaine en quadrilatères (en dimension 2) et en hexaèdres (en dimension 3).

La forme géométrique des éléments étant choisie, on suppose que nous avons un maillage comportant  $nel$  éléments tel qu'illustré à la figure 6.10.

### 6.2.1 Les noeuds

Sur chaque élément  $K$ , on définit  $n_g^K$  *noeuds géométriques*. Très souvent, les noeuds géométriques correspondent aux sommets du triangle, du quadrilatère ou du tétraèdre. Le choix des noeuds géométriques déterminera la transformation à l'élément de référence.

Comme en dimension 1, on introduit également les  $n_C$  *noeuds de calcul* où on évaluera les valeurs des variables essentielles du problème. On supposera ici encore, bien que ce ne soit pas absolument nécessaire, que les noeuds de calcul comprennent les noeuds géométriques. Le nombre total de noeuds de l'élément  $K$  est noté  $n_c^K$ .

On place les coordonnées de tous les noeuds (géométriques et/ou de calcul) du maillage dans le tableau *coor* de dimension  $nnoeuds$  multipliée par la dimension d'espace  $d$  ( $d = 2$  ou  $d$ ). Enfin, le tableau de connectivité des noeuds *connec* contient, pour chaque élément  $K$  les numéros de ses  $n_c^K$  noeuds de calcul (en plaçant les noeuds géométriques en premier).

### 6.2.2 Les degrés de liberté

On associe à chaque noeud de calcul un certain nombre de degrés de liberté (au total  $n_d^K$  sur chaque élément), dépendant du nombre de variables essentielles du problème. Ici encore, on numérotera en dernier les degrés de liberté qui sont imposés par les conditions aux limites essentielles du problème donné. On a donc la partition :

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{nddl} \end{bmatrix} = \begin{bmatrix} U^I \\ U^C \end{bmatrix}$$

où  $nddl$  désigne le nombre total de degrés de liberté du maillage.

Rappelons qu'on associe à chaque degré de liberté une fonction de Ritz qui sera construite élément par élément. La méthode de Ritz-Galerkin nous mènera à la résolution d'un système linéaire global de la forme :

$$A\delta_U = F + S \quad (6.2)$$

suivant une notation similaire à celle des éléments unidimensionnels (voir la relation 5.19).

### 6.2.3 Numérotation des degrés de liberté

S'il est relativement simple de numérotter les degrés de liberté en dimension 1, il faut être beaucoup plus prudent en dimension supérieure. En effet, cette numérotation a une influence majeure sur la structure de la matrice globale qui sera assemblée. Nous avons vu que les matrices engendrées par la méthode des éléments finis sont creuses c.-à-d. contiennent une part importante de zéros. Cette structure dépend directement du tableau *adres* qui détermine si un degré de liberté  $i$  est connecté ou non à un autre degré de liberté  $j$ . Si tel est le cas, le coefficient  $a_{ij}$  de la matrice globale sera non nul. Sinon, on a tout simplement  $a_{ij} = 0$ . Pour que 2 degrés de liberté soient connectés, il faut et il suffit qu'ils apparaissent tous deux dans une même ligne du tableau *adres*. C'est une autre façon de dire que les supports des fonctions de Ritz associées à ces degrés de liberté s'intersectent.

Introduisons immédiatement un peu de terminologie. On définit la *largeur de bande* de la ligne  $i$  du système 6.2 par la relation :

$$l_i = \max_{j \geq i | a_{ij} \neq 0} (j - i + 1)$$

Il s'agit tout simplement de la distance entre la diagonale de la matrice et le dernier terme non nul de la ligne  $i$ . En d'autres termes,  $a_{ij} = 0$  pour  $j \geq i + l_i$ . De même, la *largeur de bande maximale*  $l_{max}$  est la plus grande valeur des  $l_i$  c.-à-d. :

$$l_{max} = \max_i l_i$$

Il est important de concentrer les termes non nuls de la matrice au voisinage de la diagonale principale de façon à diminuer l'espace-mémoire requis. Cela revient à minimiser les largeurs de bande. On obtient ainsi une matrice dite *matrice bande* ou plus précisément une matrice dite en *ligne de ciel*. Cela est particulièrement important si on utilise une décomposition  $LU$  pour la résolution du système linéaire global. En effet, cette méthode préserve la structure bande de la matrice de sorte que l'on peut stocker la décomposition  $LU$  sans augmenter la mémoire nécessaire. Sur chaque ligne du système, les termes au delà d'une distance  $l_i$  de la diagonale sont nuls et resteront nuls après la décomposition  $LU$ . Il suffit pour s'en convaincre de regarder l'algorithme de la décomposition  $LU$  (voir Fortin, réf.[17]) :

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj} \quad \text{et} \quad u_{ij} = \frac{a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}}{l_{ii}}$$

En pratique, on donne une numérotation quelconque des degrés de liberté dans le tableau *numer*. Par la suite, on passe les tableaux *numer* et *convec* à un programme d'optimisation qui renumérottera les degrés de liberté (en modifiant le tableau *numer*) et créera le tableau *adres* suivant un certain critère d'optimisation. On voudra par exemple minimiser la largeur de bande maximale  $l_{max}$  de la matrice ou encore minimiser :

$$\sum_i l_i$$

qui est une indication de l'espace-mémoire nécessaire. Différents algorithmes de numérotation des degrés de liberté existent et nous ne mentionnerons que celui de Cuthill-McKee [22].

Pour illustrer ce point, on présente à la figure 6.1 un exemple typique de matrice creuse (de dimension 60 par 60) semblables à celles obtenues par éléments finis. On indique seulement les termes non nuls par un point. Si on renumérote les degrés de liberté à l'aide de l'algorithme de Cuthill-McKee, on obtient la deuxième matrice de la figure 6.1. On remarque aisément la diminution de la largeur de bande entre les deux matrices illustrées. Ainsi, entre les deux numérotations (entre les deux vecteurs *numer*), la largeur de bande maximale  $l_{max}$  est passée de 35 à 12. Il en résulte une économie d'espace-mémoire et généralement une plus grande rapidité de résolution du système linéaire.

### 6.3 Formulation variationnelle élémentaire

La formulation variationnelle élémentaire s'obtient à partir de l'équation aux dérivées partielles de départ mais en intégrant cette fois sur un élément  $K$ . On obtient ainsi à l'aide de la relation A.5 :

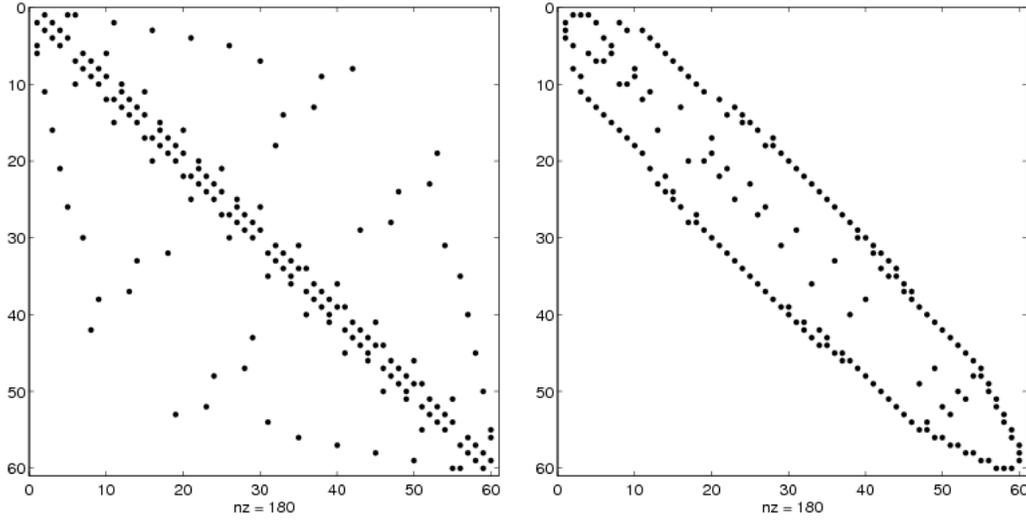


Figure 6.1 – Structure de la matrice avant et après renumérotation

$$\int_K (p(\mathbf{x})u(\mathbf{x})w(\mathbf{x}) + q(\mathbf{x})\nabla u \cdot \nabla w) dv = \int_K r(\mathbf{x})w(\mathbf{x}) dv + \int_{\partial K} s^K(\mathbf{x})w(\mathbf{x}) ds$$

On a ainsi introduit la variable secondaire :

$$s^K(\mathbf{x}) = q(\mathbf{x}) \frac{\partial u}{\partial \mathbf{n}^K}$$

où  $\mathbf{n}^K$  est le vecteur normal extérieur à la frontière  $\partial K$  de  $K$ . La variable  $s^K(\mathbf{x})$  est la condition naturelle de ce problème. On passe maintenant en correction :

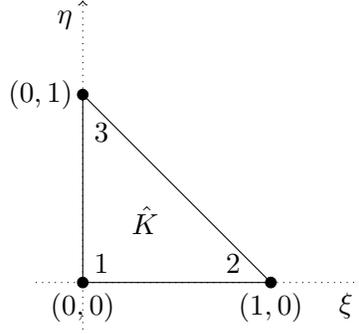
$$\begin{aligned} & \int_K (p(\mathbf{x})\delta_u(\mathbf{x})w(\mathbf{x}) + q(\mathbf{x})\nabla \delta_u \cdot \nabla w) dv \\ &= \int_K r(\mathbf{x})w(\mathbf{x}) dv - \int_K (p(\mathbf{x})u_g(\mathbf{x})w(\mathbf{x}) + q(\mathbf{x})\nabla u_g \cdot \nabla w) dv + \int_{\partial K} s^K(\mathbf{x})w(\mathbf{x}) ds \end{aligned}$$

On applique la méthode de Ritz sur chaque élément  $K$  en posant :

$$\delta_u(\mathbf{x})|_K \simeq \delta_u^K(\mathbf{x}) = \sum_{j=1}^{n_d^K} \delta_{u_j^K} \psi_j^K(\mathbf{x}) \quad \text{et} \quad u_g^K(\mathbf{x}) = \sum_{j=1}^{n_d^K} u_{g_j^K} \psi_j^K(\mathbf{x}) \quad (6.3)$$

Remplaçant dans la formulation variationnelle et en prenant successivement  $w(\mathbf{x}) = \psi_i^K(\mathbf{x})$ , pour  $i$  allant de 1 jusqu'à  $n_d^K$ . On obtient ainsi le système :

$$A^K \delta_U^K = F^K + S^K$$

Figure 6.2 – Éléments de référence triangulaire à 3 noeuds ( $n_g^K = 3$ )

où :

$$a_{ij}^K = \int_K (p(\mathbf{x})\psi_j^K(\mathbf{x})\psi_i^K(\mathbf{x}) + q(\mathbf{x})\nabla\psi_j^K(\mathbf{x}) \cdot \nabla\psi_i^K(\mathbf{x})) \, dv$$

$$f_i^K = \int_K r(\mathbf{x})\psi_i^K(\mathbf{x}) \, dv - \int_K (p(\mathbf{x})u_g^K(\mathbf{x})\psi_i^K(\mathbf{x}) + q(\mathbf{x})\nabla u_g^K(\mathbf{x}) \cdot \nabla\psi_i^K(\mathbf{x})) \, dv$$

et :

$$s_i^K = \int_{\partial K} s^K(\mathbf{x})\psi_i^K(\mathbf{x}) \, ds$$

## 6.4 Passage à l'élément de référence

C'est ici que le choix de la forme de l'élément et le choix des noeuds géométriques prennent beaucoup d'importance. Tout comme en dimension 1, l'élément de référence  $\hat{K}$  est un élément sur lequel on effectue tous les calculs nécessaires à l'obtention du système élémentaire. Ceci n'est possible qu'après un changement de variables. Contrairement au cas unidimensionnel, plusieurs choix d'éléments de référence sont envisageables et certains sont illustrés aux figures 6.2 à 6.6. Il reste à déterminer la transformation de l'élément réel à l'élément de référence que nous appellerons encore ici la *transformation géométrique*.

Pour éviter une lourdeur excessive de la notation, nous choisissons de prendre comme vecteur position  $\mathbf{x} = (x, y, z)$  au lieu de  $(x_1, x_2, x_3)$ . De même sur l'élément de référence, on prendra  $\boldsymbol{\xi} = (\xi, \eta, \zeta)$ .

À chaque noeud géométrique  $\mathbf{x}_i^K = (x_i^K, y_i^K, z_i^K)$  de l'élément  $K$  doit correspondre un noeud géométrique  $\boldsymbol{\xi}_i = (\xi_i, \eta_i, \zeta_i)$  sur l'élément de référence  $\hat{K}$ . La transformation  $T^K$  vérifie donc :

$$T^K(\boldsymbol{\xi}_i) = \mathbf{x}_i^K \quad \text{ou inversement} \quad (T^K)^{-1}(\mathbf{x}_i^K) = \boldsymbol{\xi}_i \quad i = 1, 2, 3, \dots, n_g^K$$

Pour déterminer la transformation géométrique, il suffit de trouver une base de polynômes de dimension égale au nombre de noeuds géométriques. On construit ensuite  $n_g^K$  fonctions d'interpolation

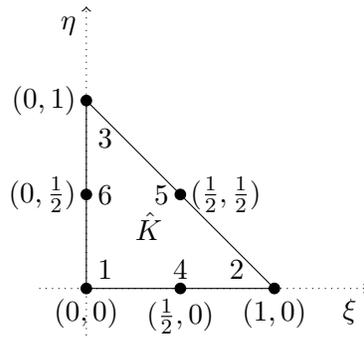


Figure 6.3 – Élément de référence triangulaire à 6 noeuds ( $n_g^K = 6$ )

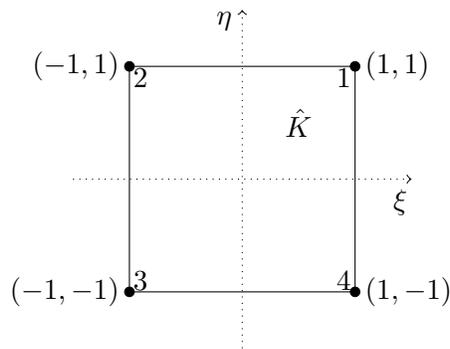


Figure 6.4 – Élément de référence quadrangulaire à 4 noeuds ( $n_g^K = 4$ )

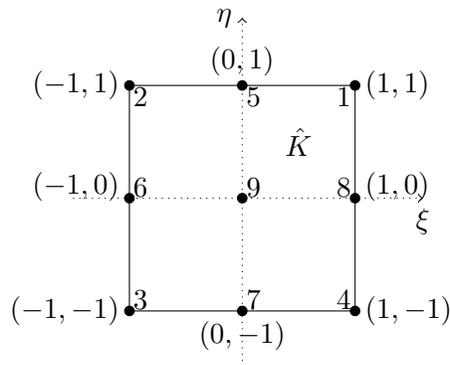
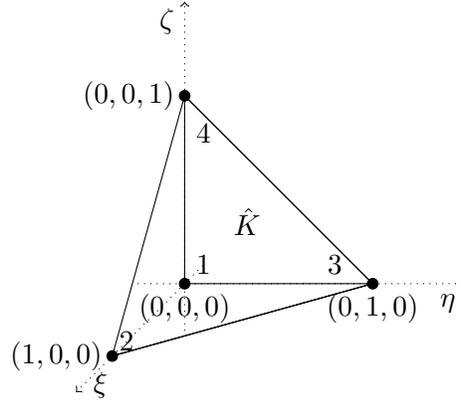


Figure 6.5 – Élément de référence quadrangulaire à 9 noeuds ( $n_g^K = 9$ )

Figure 6.6 – Élément de référence tétraédrique à 4 noeuds ( $n_g^K = 4$ )

géométriques vérifiant :

$$L_i(\boldsymbol{\xi}_j) = \delta_j^i \quad (6.4)$$

C'est la base même de l'interpolation de Lagrange (voir l'annexe C) et les fonctions  $L_i(\boldsymbol{\xi})$  sont bien sûr les fonctions d'interpolation de Lagrange. Il est facile de vérifier que la transformation :

$$T^K : \quad \hat{K} \quad \rightarrow \quad K$$

$$\boldsymbol{\xi} = (\xi, \eta, \zeta) \quad \rightarrow \quad \mathbf{x} = (x, y, z) = \sum_{i=1}^{n_g^K} \mathbf{x}_i^K L_i(\boldsymbol{\xi}) \quad (6.5)$$

possède les propriétés désirées. Il faudra ensuite transformer les dérivées des fonctions d'interpolation. C'est la technique classique de la dérivation en chaîne. Pour ce faire, il est utile d'introduire la *matrice jacobienne*  $DT^K$  associée et que l'on définit par :

$$DT^K = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \zeta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \zeta} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \zeta} \end{bmatrix}$$

Il est facile de déterminer les coefficients de cette matrice, simplement en dérivant la relation 6.5. La matrice jacobienne doit être inversible pour que la transformation  $T^K$  le soit aussi. Il suffit donc que le déterminant  $J^K$  de cette matrice soit non nul. On appelle ce déterminant le *jacobien* de la

transformation. Toute fonction d'interpolation  $\psi^K(x, y, z)$  est ainsi définie à partir de l'élément de référence par la relation :

$$\psi^K(x, y, z) = \psi^K(T^K(\xi, \eta, \zeta)) = \hat{\psi}(\xi, \eta, \zeta)$$

Pour transformer les dérivées partielles, la dérivation en chaîne nous donne :

$$\begin{bmatrix} \frac{\partial \psi^K(x, y, z)}{\partial x} \\ \frac{\partial \psi^K(x, y, z)}{\partial y} \\ \frac{\partial \psi^K(x, y, z)}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \eta}{\partial x} & \frac{\partial \zeta}{\partial x} \\ \frac{\partial \xi}{\partial y} & \frac{\partial \eta}{\partial y} & \frac{\partial \zeta}{\partial y} \\ \frac{\partial \xi}{\partial z} & \frac{\partial \eta}{\partial z} & \frac{\partial \zeta}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{\partial \hat{\psi}(\xi, \eta, \zeta)}{\partial \xi} \\ \frac{\partial \hat{\psi}(\xi, \eta, \zeta)}{\partial \eta} \\ \frac{\partial \hat{\psi}(\xi, \eta, \zeta)}{\partial \zeta} \end{bmatrix} \quad (6.6)$$

ou encore sous forme compacte :

$$[\nabla_{\mathbf{x}} \psi^K(x, y, z)] = B^K [\nabla_{\boldsymbol{\xi}} \hat{\psi}(\xi, \eta, \zeta)]$$

La matrice  $B^K$  est essentielle pour l'évaluation du système élémentaire. On l'obtient en constatant que :

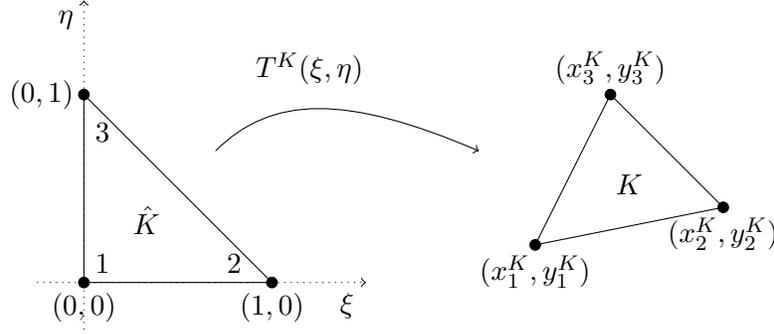
$$\begin{bmatrix} \frac{\partial \hat{\psi}(\xi, \eta, \zeta)}{\partial \xi} \\ \frac{\partial \hat{\psi}(\xi, \eta, \zeta)}{\partial \eta} \\ \frac{\partial \hat{\psi}(\xi, \eta, \zeta)}{\partial \zeta} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} & \frac{\partial z}{\partial \eta} \\ \frac{\partial x}{\partial \zeta} & \frac{\partial y}{\partial \zeta} & \frac{\partial z}{\partial \zeta} \end{bmatrix} \begin{bmatrix} \frac{\partial \psi^K(x, y, z)}{\partial x} \\ \frac{\partial \psi^K(x, y, z)}{\partial y} \\ \frac{\partial \psi^K(x, y, z)}{\partial z} \end{bmatrix} \quad (6.7)$$

qui s'écrit également :

$$[\nabla_{\boldsymbol{\xi}} \hat{\psi}(\xi, \eta, \zeta)] = (DT^K)^\top [\nabla_{\mathbf{x}} \psi^K(x, y, z)]$$

En combinant les relations 6.6 et 6.7, on conclut que la matrice de passage  $B^K$  n'est autre que  $(DT^K)^{-\top}$  c'est-à-dire la transposée de l'inverse de la matrice jacobienne. On a ainsi :

$$[\nabla_{\mathbf{x}} \psi^K(x, y, z)] = (DT^K)^{-\top} [\nabla_{\boldsymbol{\xi}} \hat{\psi}(\xi, \eta, \zeta)] \quad (6.8)$$

Figure 6.7 – Transformation linéaire sur un triangle ( $n_g^K = 3$ )

Le système élémentaire devient ensuite :

$$\begin{aligned}
 a_{ij}^K &= \int_K (p(\mathbf{x})\psi_j^K(\mathbf{x})\psi_i^K(\mathbf{x}) + q(\mathbf{x})\nabla\psi_j^K(\mathbf{x}) \cdot \nabla\psi_i^K(\mathbf{x})) \, dv \\
 &= \int_K \left( p(\mathbf{x})\psi_j^K(\mathbf{x})\psi_i^K(\mathbf{x}) + q(\mathbf{x}) [\nabla_{\mathbf{x}}\psi_j^K(x, y, z)]^\top [\nabla_{\mathbf{x}}\psi_i^K(x, y, z)] \right) \, dv \\
 &= \int_{\hat{K}} \left( p(T^K(\boldsymbol{\xi}))\hat{\psi}_j(\boldsymbol{\xi})\hat{\psi}_i(\boldsymbol{\xi}) + q(T^K(\boldsymbol{\xi})) [\nabla_{\boldsymbol{\xi}}\hat{\psi}_j(\xi, \eta, \zeta)]^\top (B^K)^\top B^K [\nabla_{\boldsymbol{\xi}}\hat{\psi}_i(\xi, \eta, \zeta)] \right) J^K \, d\hat{v} \\
 f_i^K &= \int_{\hat{K}} r(T^K(\boldsymbol{\xi}))\hat{\psi}_i(\boldsymbol{\xi}) J^K \, d\hat{v} - \int_{\hat{K}} p(T^K(\boldsymbol{\xi})) \left( \sum_{j=1}^{n_d^K} u_{g_j}^K \hat{\psi}_j(\boldsymbol{\xi}) \right) \hat{\psi}_i(\boldsymbol{\xi}) J^K \, d\hat{v} \\
 &\quad - \int_{\hat{K}} q(T^K(\boldsymbol{\xi})) \left[ \left( \sum_{j=1}^{n_d^K} u_{g_j}^K \nabla_{\boldsymbol{\xi}} \hat{\psi}_j(\xi, \eta, \zeta) \right) \right]^\top (B^K)^\top B^K [\nabla_{\boldsymbol{\xi}} \hat{\psi}_i(\xi, \eta, \zeta)] J^K \, d\hat{v} \\
 s_i^K &= \int_{\partial K} s^K(\mathbf{x})\psi_i^K(\mathbf{x}) \, ds
 \end{aligned} \tag{6.9}$$

Remarquons que l'intégrale sur la frontière des éléments n'a pas été transformée. On le fera plus tard si le besoin s'en fait sentir.

### ► Exemple 6.1

Considérons en premier lieu des transformations géométriques linéaires en dimension 2. Tout d'abord sur le triangle à 3 noeuds géométriques ( $n_g^K = 3$ ) de la figure 6.7. L'élément de référence est indiqué à la figure 6.2. La transformation  $T^K$  s'écrit à l'aide des fonctions de Lagrange C.3 sous

la forme :

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix} &= T^K(\boldsymbol{\xi}) = \sum_{i=1}^3 L_i(\boldsymbol{\xi}) \begin{bmatrix} x_i^K \\ y_i^K \end{bmatrix} = \begin{bmatrix} L_1(\boldsymbol{\xi})x_1^K + L_2(\boldsymbol{\xi})x_2^K + L_3(\boldsymbol{\xi})x_3^K \\ L_1(\boldsymbol{\xi})y_1^K + L_2(\boldsymbol{\xi})y_2^K + L_3(\boldsymbol{\xi})y_3^K \end{bmatrix} \\ &= \begin{bmatrix} (1 - \xi - \eta)x_1^K + \xi x_2^K + \eta x_3^K \\ (1 - \xi - \eta)y_1^K + \xi y_2^K + \eta y_3^K \end{bmatrix} \end{aligned}$$

Il est facile de vérifier que les fonctions  $L_i(\boldsymbol{\xi})$  vérifient la condition 6.4. La matrice jacobienne est alors :

$$DT^K = \begin{bmatrix} x_2^K - x_1^K & x_3^K - x_1^K \\ y_2^K - y_1^K & y_3^K - y_1^K \end{bmatrix}$$

Le jacobien  $J^K$  de cette transformation n'est nul que si les points  $\boldsymbol{x}_i^K$  sont colinéaires et donc si le triangle est dégénéré. De plus :

$$B^K = (DT^K)^{-\top} = \frac{1}{J^K} \begin{bmatrix} y_3^K - y_1^K & y_1^K - y_2^K \\ x_1^K - x_3^K & x_2^K - x_1^K \end{bmatrix} \quad (6.10)$$

et on montre de plus facilement que  $J^K = 2 \times \text{aire}(K)$  (en exercice). Notons enfin que le jacobien est une constante sur l'élément  $K$  et ne dépend pas de  $\boldsymbol{\xi}$ . Ce ne sera pas toujours le cas. ◀

#### ◆ Remarque 6.1

Dans le cas d'une approximation linéaire sur les éléments triangulaires, et si de plus  $p(x) = 0$ ,  $u_g(x) = 0$  et les fonctions  $q(x)$  et  $r(x)$  sont constantes par élément ( $q(x) = q^K$  et  $r(x) = r^K$  sur l'élément  $K$ ), on peut facilement évaluer le système élémentaire (voir Reddy, réf. [31]) et on obtient :

$$\begin{aligned} a_{ij}^K &= \frac{q^K}{4A^K} (\beta_i^K \beta_j^K + \gamma_i^K \gamma_j^K) \\ f_i^K &= \frac{1}{3} r^K A^K \end{aligned} \quad (6.11)$$

où  $A^K$  est l'aire de l'élément et les constantes  $\beta_i^K$  et  $\gamma_i^K$  sont données par :

$$\begin{aligned} \beta_1^K &= y_2^K - y_3^K & \gamma_1^K &= -(x_2^K - x_3^K) \\ \beta_2^K &= y_3^K - y_1^K & \gamma_2^K &= -(x_3^K - x_1^K) \\ \beta_3^K &= y_1^K - y_2^K & \gamma_3^K &= -(x_1^K - x_2^K) \end{aligned}$$

◆

#### ► Exemple 6.2

Considérons maintenant une transformation géométrique dite bilinéaire. On considère des éléments quadrangulaires à 4 sommets ( $n_g^K = 4$ ). L'élément de référence est indiqué à la figure 6.4. La transformation  $T^K$  s'écrit dans ce cas :

$$\begin{bmatrix} x \\ y \end{bmatrix} = T^K(\boldsymbol{\xi}) = \sum_{i=1}^4 L_i(\boldsymbol{\xi}) \begin{bmatrix} x_i^K \\ y_i^K \end{bmatrix} = \begin{bmatrix} L_1(\boldsymbol{\xi})x_1^K + L_2(\boldsymbol{\xi})x_2^K + L_3(\boldsymbol{\xi})x_3^K + L_4(\boldsymbol{\xi})x_4^K \\ L_1(\boldsymbol{\xi})y_1^K + L_2(\boldsymbol{\xi})y_2^K + L_3(\boldsymbol{\xi})y_3^K + L_4(\boldsymbol{\xi})y_4^K \end{bmatrix}$$

où les fonctions  $L_i(\boldsymbol{\xi})$  sont données au tableau C.7 :

$$L_1(\boldsymbol{\xi}) = \frac{1}{4}(1 + \xi)(1 + \eta)$$

$$L_2(\boldsymbol{\xi}) = \frac{1}{4}(1 - \xi)(1 + \eta)$$

$$L_3(\boldsymbol{\xi}) = \frac{1}{4}(1 - \xi)(1 - \eta)$$

$$L_4(\boldsymbol{\xi}) = \frac{1}{4}(1 + \xi)(1 - \eta)$$

La matrice jacobienne exige un peu de calcul et on obtient sous forme compacte :

$$DT^K = \begin{bmatrix} B_{11}^K + \eta C_{11}^K & B_{12}^K + \xi C_{12}^K \\ B_{21}^K + \eta C_{21}^K & B_{22}^K + \xi C_{22}^K \end{bmatrix}$$

où :

$$B_{11}^K = \frac{1}{4}(x_1^K - x_2^K - x_3^K + x_4^K) \quad C_{11}^K = \frac{1}{4}(x_1^K - x_2^K + x_3^K - x_4^K)$$

$$B_{12}^K = \frac{1}{4}(x_1^K + x_2^K - x_3^K - x_4^K) \quad C_{12}^K = \frac{1}{4}(x_1^K - x_2^K + x_3^K - x_4^K)$$

$$B_{21}^K = \frac{1}{4}(y_1^K - y_2^K - y_3^K + y_4^K) \quad C_{21}^K = \frac{1}{4}(y_1^K - y_2^K + y_3^K - y_4^K)$$

$$B_{22}^K = \frac{1}{4}(y_1^K + y_2^K - y_3^K - y_4^K) \quad C_{22}^K = \frac{1}{4}(y_1^K - y_2^K + y_3^K - y_4^K)$$

Le déterminant est de toute évidence non constant et est bien une fonction de  $\boldsymbol{\xi} = (\xi, \eta)$ . Notons de plus que :

$$B^K = (DT^K)^{-\top} = \frac{1}{J^K} \begin{bmatrix} B_{22}^K + \xi C_{22}^K & -(B_{21}^K + \eta C_{21}^K) \\ -(B_{12}^K + \xi C_{12}^K) & B_{11}^K + \eta C_{11}^K \end{bmatrix} \quad (6.12)$$

◀

### ► Exemple 6.3

Les 2 exemples précédents nécessitent l'utilisation de triangles ou de quadrilatères ayant des côtés droits. Cela est dû aux transformations linéaires (ou bilinéaires) qui assurent qu'un segment de

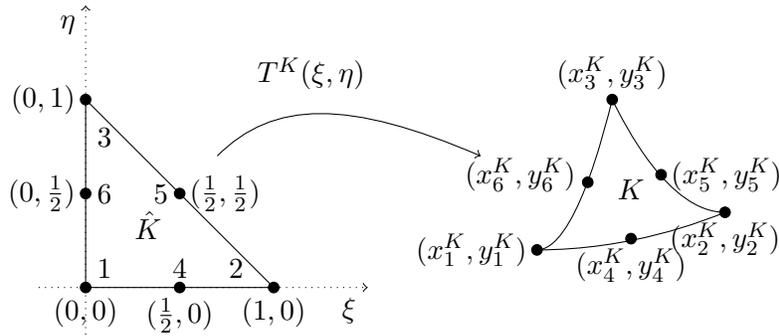


Figure 6.8 – Transformation quadratique sur un triangle ( $n_g^K = 6$ )

droite (et donc un côté droit) est transformé en un autre segment de droite. Cependant, il est parfois utile de prendre des éléments ayant des côtés courbes. Pensons par exemple à une géométrie où il y a un arc de cercle. Il est de toute évidence plus facile d’approcher un arc de cercle par des triangles (ou des quadrilatères) avec des côtés courbes.

Pour illustrer une telle situation, considérons maintenant une transformation géométrique quadratique en dimension 2 sur le triangle à 6 noeuds géométriques ( $n_g^K = 6$ ) (illustré à la figure 6.3). Cette transformation est illustrée à la figure 6.8. La transformation  $T^K$  s’écrit alors :

$$\begin{bmatrix} x \\ y \end{bmatrix} = T^K(\boldsymbol{\xi}) = \sum_{i=1}^6 L_i(\boldsymbol{\xi}) \begin{bmatrix} x_i^K \\ y_i^K \end{bmatrix}$$

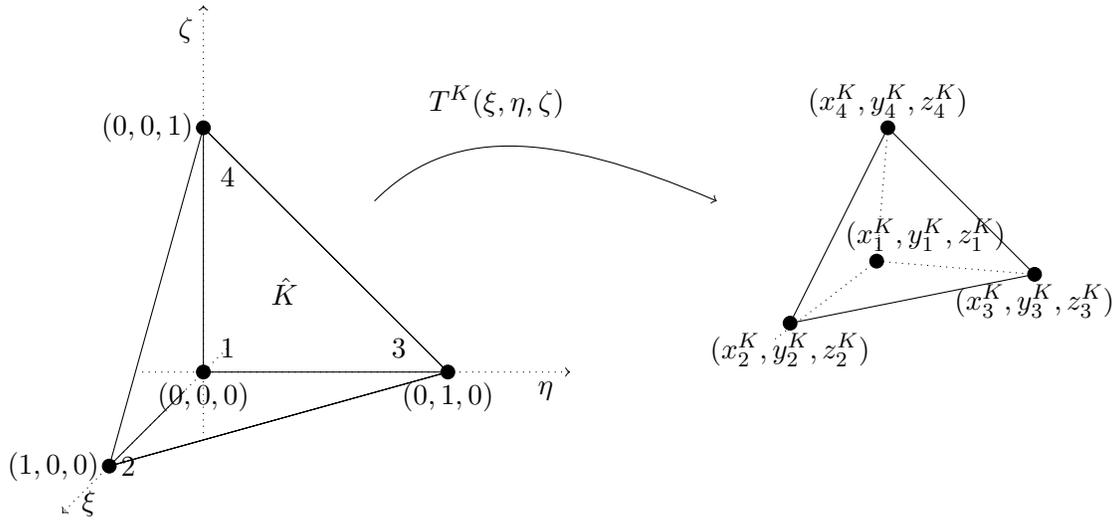
On trouvera la liste des fonctions d’interpolation  $L_i(\boldsymbol{\xi})$  au tableau C.4. ◀

► Exemple 6.4

Passons au cas tridimensionnel et considérons une transformation géométrique linéaire. Tout d’abord sur le tétraèdre à 4 noeuds géométriques ( $n_g^K = 4$ ). L’élément de référence est indiqué à la figure 6.6.

La transformation  $T^K$  (voir la figure 6.9) s’écrit alors :

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = T^K(\boldsymbol{\xi}) &= \begin{bmatrix} L_1(\boldsymbol{\xi})x_1^K + L_2(\boldsymbol{\xi})x_2^K + L_3(\boldsymbol{\xi})x_3^K + L_4(\boldsymbol{\xi})x_4^K \\ L_1(\boldsymbol{\xi})y_1^K + L_2(\boldsymbol{\xi})y_2^K + L_3(\boldsymbol{\xi})y_3^K + L_4(\boldsymbol{\xi})y_4^K \\ L_1(\boldsymbol{\xi})z_1^K + L_2(\boldsymbol{\xi})z_2^K + L_3(\boldsymbol{\xi})z_3^K + L_4(\boldsymbol{\xi})z_4^K \end{bmatrix} \\ &= \begin{bmatrix} (1 - \xi - \eta - \zeta)x_1^K + \xi x_2^K + \eta x_3^K + \zeta x_4^K \\ (1 - \xi - \eta - \zeta)y_1^K + \xi y_2^K + \eta y_3^K + \zeta y_4^K \\ (1 - \xi - \eta - \zeta)z_1^K + \xi z_2^K + \eta z_3^K + \zeta z_4^K \end{bmatrix} \end{aligned}$$

Figure 6.9 – Transformation linéaire sur un tétraèdre ( $n_g^K = 4$ )

Il est facile de vérifier que les fonctions  $L_i(\boldsymbol{\xi})$  vérifient la condition 6.4. La matrice jacobienne est alors :

$$DT^K = \begin{bmatrix} x_2^K - x_1^K & x_3^K - x_1^K & x_4^K - x_1^K \\ y_2^K - y_1^K & y_3^K - y_1^K & y_4^K - y_1^K \\ z_2^K - z_1^K & z_3^K - z_1^K & z_4^K - z_1^K \end{bmatrix}$$

Le jacobien  $J^K$  de cette transformation n'est nul que si les points  $\mathbf{x}_i^K$  sont coplanaires et donc si le tétraèdre est dégénéré. Notons de plus que si on dénote  $t_{ij}^K$  les coefficients de la matrice  $DT^K$ , la matrice inverse s'écrit :

$$B^K = (DT^K)^{-\top} = \frac{1}{J^K} \begin{bmatrix} t_{22}^K t_{33}^K - t_{32}^K t_{23}^K & t_{13}^K t_{32}^K - t_{12}^K t_{33}^K & t_{12}^K t_{23}^K - t_{13}^K t_{22}^K \\ t_{31}^K t_{23}^K - t_{21}^K t_{33}^K & t_{11}^K t_{33}^K - t_{13}^K t_{31}^K & t_{21}^K t_{13}^K - t_{23}^K t_{11}^K \\ t_{21}^K t_{32}^K - t_{31}^K t_{22}^K & t_{12}^K t_{31}^K - t_{32}^K t_{11}^K & t_{11}^K t_{22}^K - t_{12}^K t_{21}^K \end{bmatrix}$$

On montre de plus facilement que  $J^K = 6 \times \text{volume}(K)$  (en exercice). Notons qu'encore ici que le jacobien est une constante et ne dépend pas de  $\boldsymbol{\xi}$ . ◀

## 6.5 Construction des fonctions d'interpolation

Tout comme en dimension 1, à chaque noeud de calcul  $(\mathbf{x}_1^K, \mathbf{x}_2^K, \dots, \mathbf{x}_{n_C}^K)$  de l'élément réel  $K$ , correspond un noeud d'interpolation  $(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_{n_C})$  sur l'élément de référence par la relation :

$$\boldsymbol{\xi}_i = (T^K)^{-1}(\mathbf{x}_i^K) \text{ ou encore } \mathbf{x}_i^K = T^K(\boldsymbol{\xi}_i), \quad i = 1, 2, \dots, n_C$$

Sur l'élément  $K$ , à chaque degré de liberté de chaque noeud de calcul correspond une fonction d'interpolation  $\psi_j^K(\mathbf{x})$ . Ces fonctions ne seront construites que sur l'élément de référence et en servant de la transformation  $T^K$ , on aura :

$$\psi_j^K(\mathbf{x}) = \hat{\psi}_j(\boldsymbol{\xi})$$

et en particulier :

$$\psi_j^K(\mathbf{x}_i^K) = \hat{\psi}_j(\boldsymbol{\xi}_i)$$

Généralement, il n'y a nul besoin de construire explicitement les fonctions  $\psi_j^K(\mathbf{x})$  puisque nous ne travaillerons que sur l'élément de référence. La seule différence existant avec le cas unidimensionnel est la plus grande variété de transformations  $T^K$  que l'on peut utiliser.

Le choix des fonctions  $\psi_j^K(\mathbf{x})$  dépend du problème que l'on souhaite résoudre et plus particulièrement de l'espace  $V$  de la formulation variationnelle. Pour le moment nous nous limiterons aux équations aux dérivées partielles d'ordre 2 de sorte que l'espace  $V$  sera  $H^1(\Omega)$  ou l'un de ses sous-espaces. Pour construire des fonctions d'interpolation de  $H^1(\Omega)$ , rappelons que l'on doit s'assurer de la continuité à la frontière des éléments. Cela est une conséquence du théorème 2.12.

Pour les problèmes d'ordre 2, on peut utiliser l'interpolation de Lagrange décrite à la section C.2 de l'annexe C. On préfère généralement les approximations linéaires ou quadratiques, que ce soit sur les triangles, les quadrilatères, les tétraèdres ou les hexaèdres. Bien entendu, rien n'empêche d'utiliser des polynômes de degré plus élevé mais le nombre de degrés de liberté augmente rapidement avec le degré des polynômes.

## 6.6 Évaluation du système élémentaire

Pour évaluer les coefficients du système élémentaire 6.9, on recourt le plus souvent à l'intégration numérique, bien que l'intégration exacte puisse être utile dans les cas très simples (interpolation de bas degré et propriétés  $p(x)$ ,  $q(x)$  et  $r(x)$  constantes par exemple).

Il n'y a pas de difficultés particulières mis-à-part le fait que le nombre de points d'intégration augmente lorsqu'on passe en dimension 2 ou 3. Ainsi, on trouvera à l'annexe D les points et les poids d'intégration de quelques unes des quadratures les plus utilisées en pratique. Ainsi sur les quadrilatères et les hexaèdres, on peut utiliser les mêmes quadratures qu'en dimension 1 (voir le tableau D.1 par le biais des relations D.4 et D.5).

Sur les triangles ou les tétraèdres, on utilise les quadratures dites de Hammer données aux tableaux D.2 et D.3. Rappelons que le degré de précision de la quadrature est un critère important dans le choix de la quadrature appropriée au calcul de termes du système élémentaire.

## 6.7 Assemblage

L'assemblage suit exactement les mêmes étapes qu'en dimension 1. Bien sûr la taille des matrices élémentaires augmente avec la dimension d'espace et le degré des fonctions d'interpolation mais les principes généraux demeurent les mêmes. On utilise l'algorithme de la page 104 qui ne change nullement avec la dimension d'espace.

La fonction de Ritz  $\phi_i(\mathbf{x})$  associée au  $i^{\text{ième}}$  degré de liberté du problème résulte de l'assemblage des fonctions d'interpolation locales  $\psi_j^K(\mathbf{x})$  sur les éléments  $K$  qui contiennent le degré de liberté numéro  $i$  dans leur tableau d'adressage. On aura donc pour ces éléments :

$$\text{adres}(K, j) = i$$

pour un certain  $j$  compris entre 1 et  $n_d^K$ .

## 6.8 Imposition des conditions aux limites

Une fois tous les systèmes élémentaires assemblés et en vertu de la convention utilisée pour le tableau *numer*, on obtient un système global de la forme :

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \delta_U^I \\ 0 \end{pmatrix} = \begin{pmatrix} F_1^C \\ F_2^C \end{pmatrix} + \begin{pmatrix} S^C \\ S^I \end{pmatrix} \quad (6.13)$$

La partition de la matrice  $A$  suit directement celle du vecteur global des degrés de liberté  $U$ . Les matrices  $M_{11}$  et  $M_{22}$  sont toujours carrées et les matrices  $M_{12}$  et  $M_{21}$  sont rectangulaires. Si la forme bilinéaire du problème est symétrique, on a  $M_{12}^T = M_{21}$ . Les matrices  $M_{12}$  et  $M_{22}$  pourraient ne pas être assemblées puisqu'elles n'ont aucun rôle par la suite.

Enfin, il reste à analyser le terme de droite composé de 2 parties. Le vecteur  $F$  est entièrement déterminé et ne pose aucun problème. Par contre, le vecteur  $S$  contenant la contribution des variables secondaires est lui aussi décomposé en 2 parties. Là où la variable essentielle est imposée (et donc connue), nous avons vu que la condition naturelle est inconnue et vice versa.

Bien que similaire, la situation est légèrement plus complexe que dans le cas unidimensionnel et mérite donc une attention particulière. Rappelons que le vecteur  $S$  est constitué de l'assemblage de termes de la forme :

$$s_i^K = \int_{\partial K} s^K(\mathbf{x}) \psi_i^K(\mathbf{x}) ds$$

où :

$$s^K(\mathbf{x}) = q(\mathbf{x}) \frac{\partial u}{\partial \mathbf{n}^K}$$

Comme nous le verrons plus loin, le traitement du vecteur  $S$  est sensiblement plus délicat que dans le cas unidimensionnel.

## 6.9 Résolution du système linéaire global

Pour la résolution du système linéaire on procède toujours en 2 étapes. Tout d'abord, on détermine le vecteur  $U^I$  en résolvant le système linéaire :

$$M_{11}\delta_U^I = F_1^C + S^C \quad (6.14)$$

qui n'est qu'une réécriture de la première équation du système 6.13. On remarque que le terme de droite est entièrement connu. Pour cela, on utilise les techniques classiques comme la méthode de décomposition  $LU$ . Remarquons que cette équation n'est rien d'autre que la discrétisation de la forme variationnelle :

$$a(\delta_u, w) = l(w) - a(u_g, w)$$

puisque le vecteur  $U^C$  correspond au relèvement des conditions essentielles  $u_g$ .

Une fois le vecteur  $U^I$  calculé, on détermine si nécessaire le vecteur  $S^I$  directement en posant :

$$S^I = M_{21}\delta_U^I - F_2^C \quad (6.15)$$

## 6.10 Présentation des résultats

Nous n'insisterons jamais assez sur l'importance d'un bon visualisateur, particulièrement en dimension 2 et 3. Si en dimension 1 on peut se débrouiller avec des instruments graphiques simples, ce n'est plus le cas en dimension supérieure. Pour les résultats qui suivent, nous utiliserons le logiciel VU développé par Benoît Ozell [27] et qui possède toutes les fonctionnalités requises en dimension 2 ou 3.

Un bon logiciel de visualisation sera capable de produire rapidement des courbes d'isovaleurs des différentes variables calculées, des champs de vecteurs, des coupes de toutes sortes, qui permettent de donner une signification à des colonnes de chiffres qui autrement seraient difficilement utilisables.

## 6.11 Exemples et applications

Dans cette section, nous présentons quelques exemples simples sur des maillages de petite taille pour illustrer la méthode et les différentes étapes à suivre. Nous terminons par des applications dans différents domaines.

### ► Exemple 6.5

Nous pouvons dès maintenant passer à un exemple relativement simple. La géométrie du domaine est le carré  $[0, 1]^2$ . On résoudra l'équation de Poisson :

$$-\nabla^2 u(x) = 1$$

avec des conditions essentielles homogènes sur toute la frontière ( $u_g(x) = 0$ ).

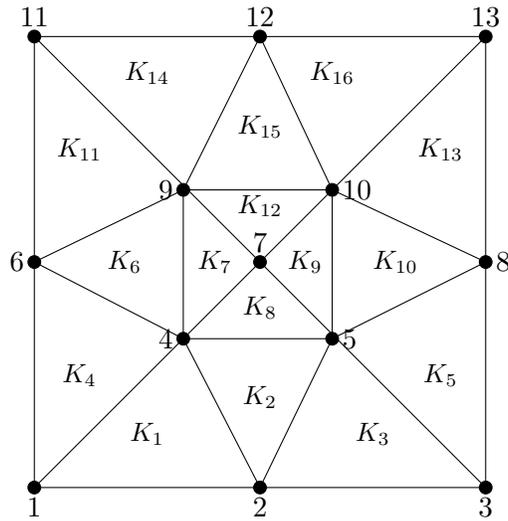


Figure 6.10 – Numérotation des éléments et des noeuds

– **Le maillage**

On utilise le maillage de la figure 6.10 constitué de 16 éléments et de 13 noeuds. Nous choisissons une interpolation linéaire sur chaque élément ( $n_C = n_g^K = 3$ ). Le tableau *coor* prend la forme :

Coordonnées des noeuds					
Tableau <i>coor</i>					
Noeud	Composante $x_1$	Composante $x_2$	Noeud	Composante $x_1$	Composante $x_2$
1	0,0000	0,0000	8	1,0000	0,5000
2	0,5000	0,0000	9	0,3333	0,6667
3	1,0000	0,0000	10	0,6667	0,6667
4	0,3333	0,3333	11	0,0000	1,0000
5	0,6667	0,3333	12	0,5000	1,0000
6	0,0000	0,5000	13	1,0000	1,0000
7	0,5000	0,5000			

Numéros des noeuds							
Tableau <i>connec</i>							
	Noeuds				Noeuds		
Élément	#1	#2	#3	Élément	#1	#2	#3
1	1	2	4	9	5	10	7
2	2	5	4	10	8	10	5
3	2	3	5	11	11	6	9
4	6	1	4	12	10	9	7
5	3	8	5	13	8	13	10
6	6	4	9	14	12	11	9
7	9	4	7	15	12	9	10
8	4	5	7	16	13	12	10

On construit le tableau *numer* en 2 étapes en évitant dans un premier temps de numéroter les degrés de liberté qui sont fixés par les conditions aux limites essentielles. On obtient ainsi :

Numérotation		Numérotation	
Tableau <i>numer</i>		Tableau <i>numer</i>	
Noeud	<i>u</i>	Noeud	<i>u</i>
1	?	1	6
2	?	2	7
3	?	3	8
4	1	4	1
5	2	5	2
6	?	6	9
7	3	7	3
8	?	8	10
9	4	9	4
10	5	10	5
11	?	11	11
12	?	12	12
13	?	13	13

Enfin, on peut déterminer le tableau d'adressage *adres* à l'aide des tableaux de connectivité *connec* et de numérotation *numer* (voir les exercices de fin de chapitre) :

Numéros des ddl							
Tableau <i>adres</i>							
Élément	Ddl #1	Ddl #2	Ddl #3	Élément	Ddl #1	Ddl #2	Ddl #3
1	6	7	1	9	2	5	3
2	7	2	1	10	10	5	2
3	7	8	2	11	11	9	4
4	9	6	1	12	5	4	3
5	8	10	2	13	10	13	5
6	9	1	4	14	12	11	4
7	4	1	3	15	12	4	5
8	1	2	3	16	13	12	5

– **Les systèmes élémentaires**

Il reste à construire les systèmes élémentaires. On peut directement utiliser la relation 6.11 puisque dans cet exemple,  $p(\mathbf{x}) = 0$ ,  $u_g(x) = 0$  et  $q^K = f^K = 1$  pour tous les éléments  $K$ . Pour le premier élément, on a :

$$\frac{1}{12} \begin{bmatrix} 5 & -2 & -3 \\ -2 & 8 & -6 \\ -3 & -6 & 9 \end{bmatrix} \begin{bmatrix} \delta_{u_1}^{K_1} \\ \delta_{u_2}^{K_1} \\ \delta_{u_3}^{K_1} \end{bmatrix} = \frac{1}{36} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} s_1^{K_1} \\ s_2^{K_1} \\ s_3^{K_1} \end{bmatrix}$$

La numérotation des noeuds (et donc des degrés de liberté) fait en sorte qu'on obtient exactement le même système élémentaire pour les éléments 5, 11, et 16.

Pour le deuxième élément, on a :

$$\frac{1}{8} \begin{bmatrix} 4 & -2 & -2 \\ -2 & 5 & -3 \\ -2 & -3 & 5 \end{bmatrix} \begin{bmatrix} \delta_{u_1}^{K_2} \\ \delta_{u_2}^{K_2} \\ \delta_{u_3}^{K_2} \end{bmatrix} = \frac{1}{54} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} s_1^{K_2} \\ s_2^{K_2} \\ s_3^{K_2} \end{bmatrix}$$

Ici encore, le même système élémentaire est aussi valable pour les éléments 6, 10 et 15.

Pour le troisième élément (ainsi que pour les éléments 4, 13 et 14), on a :

$$\frac{1}{12} \begin{bmatrix} 8 & -2 & -6 \\ -2 & 5 & -3 \\ -6 & -3 & 9 \end{bmatrix} \begin{bmatrix} \delta_{u_1}^{K_3} \\ \delta_{u_2}^{K_3} \\ \delta_{u_3}^{K_3} \end{bmatrix} = \frac{1}{36} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} s_1^{K_3} \\ s_2^{K_3} \\ s_3^{K_3} \end{bmatrix}$$

Pour le septième élément (ainsi que pour les éléments 8, 9 et 12), on a :

$$\frac{1}{4} \begin{bmatrix} 2 & 0 & -2 \\ 0 & 2 & -2 \\ -2 & -2 & 4 \end{bmatrix} \begin{bmatrix} \delta_{u_1}^{K_7} \\ \delta_{u_2}^{K_7} \\ \delta_{u_3}^{K_7} \end{bmatrix} = \frac{1}{108} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} s_1^{K_7} \\ s_2^{K_7} \\ s_3^{K_7} \end{bmatrix}$$

– **L’assemblage**

Un système linéaire de dimension 13 par 13 doit être assemblé à partir des systèmes élémentaires que nous venons d’obtenir. On utilise exactement la même technique qu’en dimension 1 et nous nous limiterons à en rappeler brièvement les principales étapes. À l’aide du tableau *adres*, on écrit le système élémentaire du premier élément sous la forme :

$$\frac{1}{12} \left[ \begin{array}{c|ccc} & 6 & 7 & 1 \\ \hline 6 & 5 & -2 & -3 \\ 7 & -2 & 8 & -6 \\ 1 & -3 & -6 & 9 \end{array} \right] \begin{bmatrix} \delta_{u_1}^{K_1} \\ \delta_{u_2}^{K_1} \\ \delta_{u_3}^{K_1} \end{bmatrix} = \frac{1}{36} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} s_1^{K_1} \\ s_2^{K_1} \\ s_3^{K_1} \end{bmatrix}$$

On fait de même pour tous les éléments et on obtient (en exercice) :

$$\frac{1}{24} \left[ \begin{array}{ccccc|ccccc|ccccc} 90 & -9 & -24 & -9 & 0 & -12 & -18 & 0 & -18 & 0 & 0 & 0 & 0 \\ -9 & 90 & -24 & 0 & -9 & 0 & -18 & -12 & 0 & -18 & 0 & 0 & 0 \\ -24 & -24 & 96 & -24 & -24 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -9 & 0 & -24 & 90 & -9 & 0 & 0 & 0 & -18 & 0 & -12 & -18 & 0 \\ 0 & -9 & -24 & -9 & 90 & 0 & 0 & 0 & 0 & -18 & 0 & -18 & -12 \\ \hline -12 & 0 & 0 & 0 & 0 & 20 & -4 & 0 & -4 & 0 & 0 & 0 & 0 \\ -18 & -18 & 0 & 0 & 0 & -4 & 44 & -4 & 0 & 0 & 0 & 0 & 0 \\ 0 & -12 & 0 & 0 & 0 & 0 & -4 & 20 & 0 & -4 & 0 & 0 & 0 \\ -18 & 0 & 0 & -18 & 0 & -4 & 0 & 0 & 44 & 0 & -4 & 0 & 0 \\ 0 & -18 & 0 & 0 & -18 & 0 & 0 & -4 & 0 & 44 & 0 & 0 & -4 \\ 0 & 0 & 0 & -12 & 0 & 0 & 0 & 0 & -4 & 0 & 20 & -4 & 0 \\ 0 & 0 & 0 & -18 & -18 & 0 & 0 & 0 & 0 & 0 & -4 & 44 & -4 \\ 0 & 0 & 0 & 0 & -12 & 0 & 0 & 0 & 0 & -4 & 0 & -4 & 20 \end{array} \right] \begin{bmatrix} \delta_{u_1} \\ \delta_{u_2} \\ \delta_{u_3} \\ \delta_{u_4} \\ \delta_{u_5} \\ \delta_{u_6} \\ \delta_{u_7} \\ \delta_{u_8} \\ \delta_{u_9} \\ \delta_{u_{10}} \\ \delta_{u_{11}} \\ \delta_{u_{12}} \\ \delta_{u_{13}} \end{bmatrix} = \frac{1}{54} \begin{bmatrix} 6 \\ 6 \\ 2 \\ 6 \\ 6 \\ \hline 3 \\ 4 \\ 3 \\ 4 \\ 4 \\ 3 \\ 4 \\ 3 \end{bmatrix} + \begin{bmatrix} s_3^{K_1} + s_3^{K_4} + s_3^{K_2} + s_2^{K_6} + s_2^{K_7} + s_1^{K_8} \\ s_3^{K_3} + s_3^{K_5} + s_2^{K_2} + s_3^{K_{10}} + s_2^{K_8} + s_1^{K_9} \\ s_3^{K_7} + s_3^{K_8} + s_3^{K_9} + s_3^{K_{12}} \\ s_3^{K_{11}} + s_3^{K_{14}} + s_3^{K_6} + s_2^{K_{15}} + s_1^{K_7} + s_2^{K_{12}} \\ s_3^{K_{13}} + s_3^{K_{16}} + s_2^{K_{10}} + s_3^{K_{15}} + s_2^{K_9} + s_1^{K_{12}} \\ \hline s_1^{K_1} + s_2^{K_4} \\ s_2^{K_1} + s_1^{K_3} + s_1^{K_2} \\ s_2^{K_3} + s_1^{K_5} \\ s_1^{K_4} + s_2^{K_{11}} + s_1^{K_6} \\ s_2^{K_5} + s_1^{K_{13}} + s_1^{K_{10}} \\ s_1^{K_{11}} + s_2^{K_{14}} \\ s_1^{K_{14}} + s_2^{K_{16}} + s_1^{K_{15}} \\ s_2^{K_{13}} + s_1^{K_{16}} \end{bmatrix}$$

qui est encore ici de la forme de la partition 5.19.

– **Imposition des conditions aux limites**

Pour ce problème, les conditions aux limites essentielles sont homogènes (nulles). Les inconnues  $u_6$  jusqu’à  $u_{13}$  prennent donc la valeur 0.

Il reste à considérer le vecteur  $S$  des conditions naturelles. En principe, le vecteur  $S^C$  est connu mais cela ne semble pas évident lorsqu'on regarde le système global obtenu. Il faut donc y regarder de plus près. Considérons donc l'expression :

$$s_1 = s_3^{K_1} + s_3^{K_4} + s_3^{K_2} + s_2^{K_6} + s_2^{K_7} + s_1^{K_8}$$

Le raisonnement qui suit pourra s'appliquer aux termes  $s_2$  jusqu'à  $s_5$ . On a :

$$s_3^{K_1} = \int_{\partial K_1} s^{K_1}(\mathbf{x}) \psi_3^{K_1}(\mathbf{x}) ds$$

en rappelant que puisque  $q(\mathbf{x}) = 1$  :

$$s^{K_1} = \frac{\partial u}{\partial \mathbf{n}^{K_1}}$$

On peut alors écrire :

$$s_3^{K_1} = \int_{C_1^{K_1}} s^{K_1} \psi_3^{K_1} ds + \int_{C_2^{K_1}} s^{K_1} \psi_3^{K_1} ds + \int_{C_3^{K_1}} s^{K_1} \psi_3^{K_1} ds$$

où  $C_i^{K_1}$  désigne le  $i^{\text{ième}}$  côté de l'élément  $K_1$ . La fonction  $\psi_3^{K_1}$  est linéaire et prend la valeur 1 sur le troisième sommet de l'élément  $K_1$  et s'annule sur les 2 autres sommets. Elle s'annule donc identiquement sur le côté  $C_1^{K_1}$  entre les sommets 1 et 2 ce qui permet d'éliminer le premier terme de cette dernière expression. Il en est de même pour les autres coefficients et on a :

$$\begin{aligned} s_1 &= \int_{C_2^{K_1}} s^{K_1} \psi_3^{K_1} ds + \int_{C_3^{K_1}} s^{K_1} \psi_3^{K_1} ds + \int_{C_2^{K_4}} s^{K_4} \psi_3^{K_4} ds + \int_{C_3^{K_4}} s^{K_4} \psi_3^{K_4} ds + \\ &\int_{C_2^{K_2}} s^{K_2} \psi_3^{K_2} ds + \int_{C_3^{K_2}} s^{K_2} \psi_3^{K_2} ds + \int_{C_1^{K_6}} s^{K_6} \psi_2^{K_6} ds + \int_{C_2^{K_6}} s^{K_6} \psi_2^{K_6} ds + \\ &\int_{C_1^{K_7}} s^{K_7} \psi_2^{K_7} ds + \int_{C_2^{K_7}} s^{K_7} \psi_2^{K_7} ds + \int_{C_1^{K_8}} s^{K_8} \psi_1^{K_8} ds + \int_{C_3^{K_8}} s^{K_8} \psi_1^{K_8} ds \end{aligned}$$

que l'on regroupe de la manière suivante :

$$\begin{aligned} s_1 &= \int_{C_2^{K_1}} s^{K_1} \psi_3^{K_1} ds + \int_{C_3^{K_2}} s^{K_2} \psi_3^{K_2} ds + \int_{C_3^{K_1}} s^{K_1} \psi_3^{K_1} ds + \int_{C_2^{K_4}} s^{K_4} \psi_3^{K_4} ds + \\ &\int_{C_3^{K_4}} s^{K_4} \psi_3^{K_4} ds + \int_{C_1^{K_6}} s^{K_6} \psi_2^{K_6} ds + \int_{C_2^{K_2}} s^{K_2} \psi_3^{K_2} ds + \int_{C_1^{K_8}} s^{K_8} \psi_1^{K_8} ds + \\ &\int_{C_2^{K_6}} s^{K_6} \psi_2^{K_6} ds + \int_{C_1^{K_7}} s^{K_7} \psi_2^{K_7} ds + \int_{C_2^{K_7}} s^{K_7} \psi_2^{K_7} ds + \int_{C_3^{K_8}} s^{K_8} \psi_1^{K_8} ds \end{aligned}$$

Regardons les deux premiers termes. La simplification est immédiate si on constate à l'aide de la figure 6.10 que  $C_2^{K_1}$  et  $C_3^{K_2}$  sont en fait le même côté et que de plus, restreintes à ce côté, les fonctions de base  $\psi_3^{K_1}$  et  $\psi_3^{K_2}$  sont une seule et même fonction. On a alors :

$$\int_{C_2^{K_1}} s^{K_1} \psi_3^{K_1} ds + \int_{C_3^{K_2}} s^{K_2} \psi_3^{K_2} ds = \int_{C_2^{K_1}} (s^{K_1} + s^{K_2}) \psi_3^{K_1} ds = \int_{C_2^{K_1}} \left( \frac{\partial u}{\partial \mathbf{n}^{K_1}} + \frac{\partial u}{\partial \mathbf{n}^{K_2}} \right) \psi_3^{K_1} ds$$

Mais puisque :

$$\mathbf{n}^{K_2} = -\mathbf{n}^{K_1}$$

on obtient :

$$\int_{C_2^{K_1}} \left( \frac{\partial u}{\partial \mathbf{n}^{K_1}} - \frac{\partial u}{\partial \mathbf{n}^{K_1}} \right) \psi_3^{K_1} ds = 0$$

terme qui fait intervenir le saut de la variable secondaire à l'interface entre ces 2 éléments adjacents. En se rappelant la relation 2.19, on constate que ce saut est nul car autrement un terme source (une simple couche) serait présent à cet endroit. Il en va de même pour les 5 autres couples d'intégrales curvilignes et par conséquent, on a  $s_1 = 0$ . Le même raisonnement montrerait que les termes  $s_2$  à  $s_5$  sont également nuls (voir les exercices de fin de chapitre).

#### – Solution du système linéaire

Il résulte de tout cela un système linéaire de dimension 5 que l'on peut résoudre par décomposition  $LU$ . Pour ce faire, on résout successivement les équations 6.14 et 6.15. On obtient ainsi :

$$\delta_U^I = \begin{bmatrix} 0,060185 \\ 0,060185 \\ 0,069444 \\ 0,060185 \\ 0,060185 \end{bmatrix}$$

Une fois cette étape franchie, on peut calculer (au besoin) le vecteur  $S^I$  par la relation 6.15. Le calcul nous donne le vecteur :

$$S^I = \begin{bmatrix} -8,564805 \times 10^{-2} \\ -1,643515 \times 10^{-1} \\ -8,564805 \times 10^{-2} \\ -1,643515 \times 10^{-1} \\ -1,643515 \times 10^{-1} \\ -8,564805 \times 10^{-2} \\ -1,643515 \times 10^{-1} \\ -8,564805 \times 10^{-2} \end{bmatrix}$$

Il peut être intéressant de regarder d'un peu plus près l'interprétation de ce vecteur souvent appelé *vecteur des réactions*. On vérifie sans difficulté que la somme de ce vecteur est

−0,999 998 519, la question étant de savoir pourquoi. Considérons le premier terme de ce vecteur qui s'écrit :

$$\begin{aligned}
 s_1^{K_1} + s_2^{K_4} &= \int_{C_1^{K_1}} s^{K_1} \psi_1^{K_1} ds + \int_{C_2^{K_1}} s^{K_1} \psi_1^{K_1} ds + \int_{C_3^{K_1}} s^{K_1} \psi_1^{K_1} ds + \\
 &\quad \int_{C_1^{K_4}} s^{K_4} \psi_2^{K_4} ds + \int_{C_2^{K_4}} s^{K_4} \psi_2^{K_4} ds + \int_{C_3^{K_4}} s^{K_4} \psi_2^{K_4} ds \\
 &= \int_{C_1^{K_1}} s^{K_1} \psi_1^{K_1} ds + \int_{C_2^{K_4}} s^{K_4} \psi_2^{K_4} ds \\
 &= \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} \phi_1(\mathbf{x}) ds
 \end{aligned}$$

où  $\phi_1(\mathbf{x})$  est la fonction de Ritz associée au noeud 1 du maillage c.-à-d. la fonction linéaire par élément qui vaut 1 au noeud numéro 1 et qui vaut 0 à tous les autres noeuds. Il en est de même pour toutes les autres composantes du vecteur  $S^I$  qui valent donc respectivement :

$$\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} \phi_i(\mathbf{x}) ds, \quad \text{pour } i = 1, 2, 3, 6, 8, 11, 12 \text{ et } 13$$

Il est ainsi facile de constater que lorsque l'on somme le vecteur  $S^I$ , on obtient :

$$\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} ds$$

car la somme des  $\phi_i$  est égale à 1. L'équation aux dérivées partielles de départ nous donne alors :

$$\int_{\Omega} \nabla u \cdot \nabla w dv - \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} w ds = \int_{\Omega} w dv$$

En choisissant  $w = 1$ , on trouve :

$$- \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} ds = \int_{\Omega} 1 dv = 1$$

et la somme du vecteur  $S^I$  est une très bonne approximation.

#### – Présentation des résultats

On présente les résultats à la figure 6.11. On peut y constater la simplicité de même que la symétrie de la solution. On aurait pu en fait limiter les calculs au quart de la géométrie initiale. On aurait ainsi gagné en précision.



#### ► Exemple 6.6

Nous sommes maintenant en mesure d'aborder un problème plus réaliste, mais aussi de plus grande taille. Nous ne pourrons plus expliciter les détails de tous les calculs. On considère un problème de

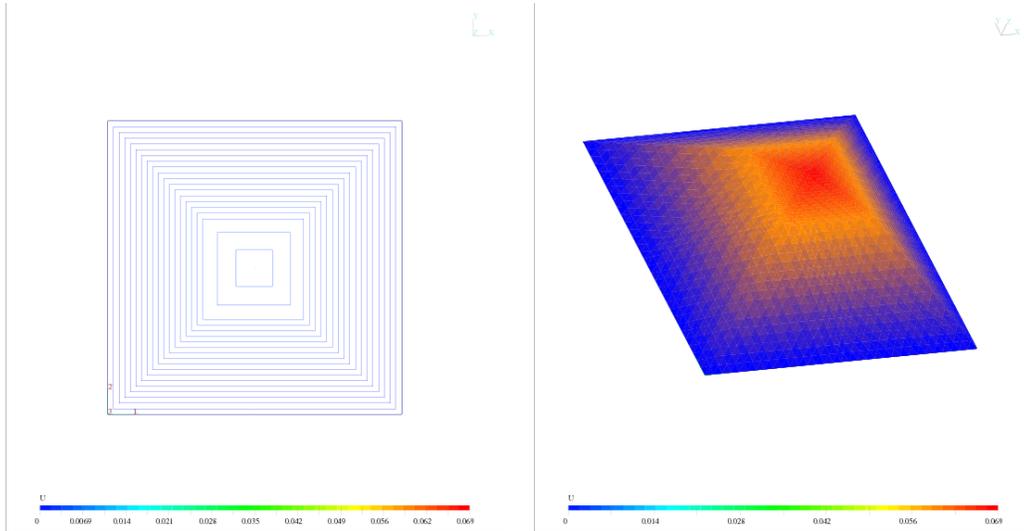


Figure 6.11 – Isovaleurs de la fonction  $u(x)$  et vue tridimensionnelle

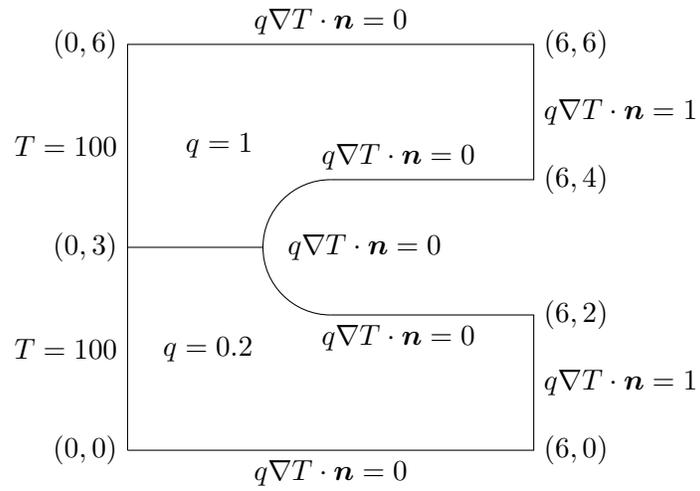


Figure 6.12 – Géométrie et conditions aux limites

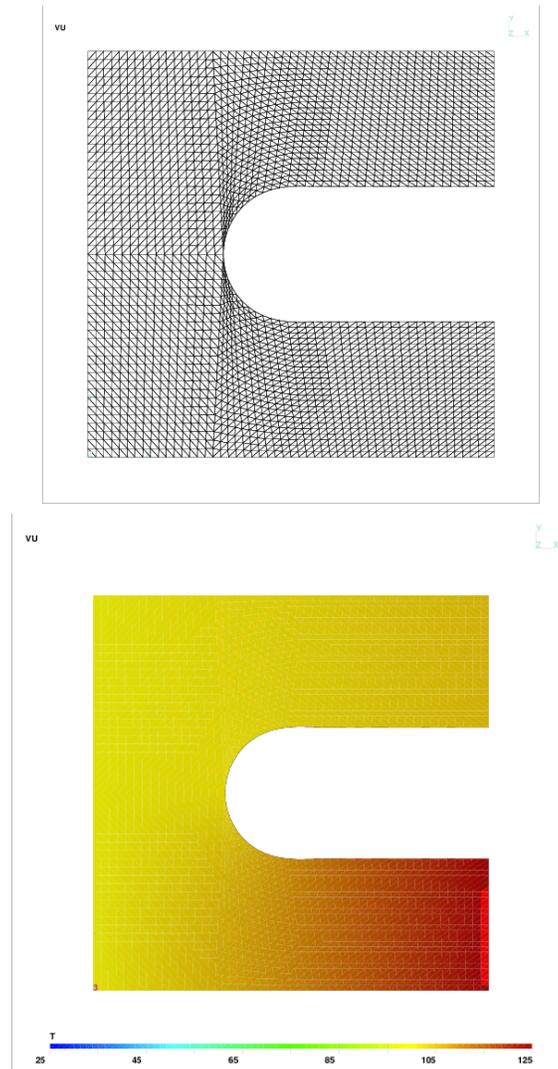


Figure 6.13 – Transfert thermique

transfert de chaleur dans une plaque métallique constituée de 2 matériaux de conductivité thermique différente tel qu'illustrée à la figure 6.12.

La partie gauche de la plaque est maintenue à une température de 100°C (condition aux limites essentielle) et elle est supposée parfaitement isolée sur les parois supérieure et inférieure de même que dans la partie en forme de U (condition aux limites naturelle nulle). Enfin, un flux de chaleur de 1 est imposée aux deux extrémités à droite. Ce problème requiert la solution de l'équation de la chaleur :

$$-\nabla \cdot (q(\mathbf{x})\nabla T) = 0$$

où  $q(\mathbf{x})$  est la conductivité thermique.

On a résolu ce problème sur le maillage de la figure 6.13. La transformation vers l'élément de référence est linéaire (voir la figure 6.7) tandis que les fonctions d'interpolation  $\hat{\psi}$  sont quadratiques (voir le tableau C.4) Le maillage est constitué de 4992 triangles (2633 sommets et 7624 arêtes) ce qui résulte en un système linéaire de 10 257 degrés de liberté.

Les isovaleurs de température sont également illustrées sur cette figure. On constate aisément la différence de comportement entre les deux moitiés du domaine. La partie de la plaque où la conductivité thermique est la plus grande (moitié supérieure) évacue la chaleur beaucoup plus facilement, ce qui évite une augmentation de la température. La température en sortie (sur l'axe  $x = 6$ ) est d'un peu moins de 126°C sur la partie inférieure et de 105,2°C sur la partie supérieure.

◀

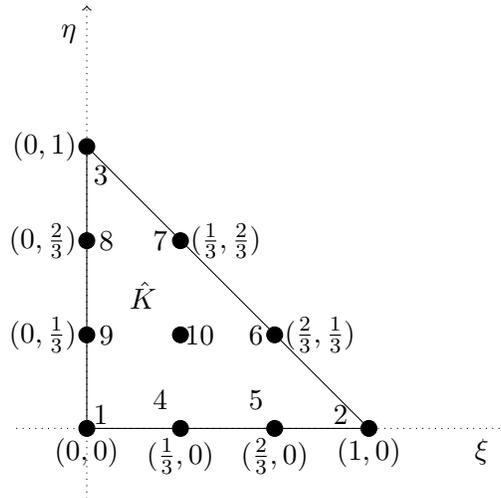


Figure 6.14 – Élément de référence triangulaire de degré 3 ( $n_g^K = 10$ )

## 6.12 Exercices

1. Obtenir la formulation variationnelle élémentaire de la page 141.
2. Vérifier les relations entre le jacobien  $J^K$  et l'aire (page 147) en dimension 2 et le volume (page 150) en dimension 3 pour les transformations linéaires de l'élément réel vers l'élément de référence.
3. Construire le tableau *adres* de la page 155.
4. Assembler le système élémentaire de la page 157.
5. Dans le terme de droite  $S$  du système de la page 157, vérifier que  $s_3 = 0$ .
6. Toujours dans le terme de droite  $S$  du système de la page 157, simplifier au maximum l'expression de  $s_6$  et  $s_7$ .
7. Obtenir les coefficients  $a_{11}^K$  et  $a_{12}^K$  dans la relation 6.11 à partir de la formulation variationnelle élémentaire 6.9.
8. Une formulation variationnelle fait apparaître le terme suivant dans un domaine bidimensionnel :

$$\int_K q(\mathbf{x}) \nabla \psi_j^K(\mathbf{x}) \cdot \nabla \psi_i^K(\mathbf{x}) \, dv$$

où  $K$  est un élément triangulaire (à côtés droits) et les  $\psi_i^K(\mathbf{x})$  sont les fonctions d'interpolation quadratiques ( $P_2$ ). Donner le nombre de points d'intégration minimal permettant d'évaluer exactement ce terme si  $q(\mathbf{x})$  est linéaire dans l'élément. Même question si cette fois on a un élément quadrangulaire (à côtés droits) et des fonctions d'interpolation  $Q_2$ .

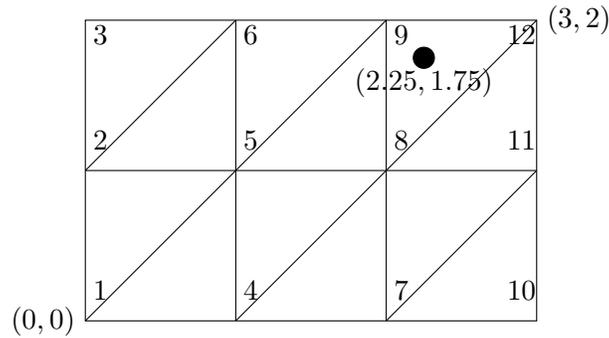
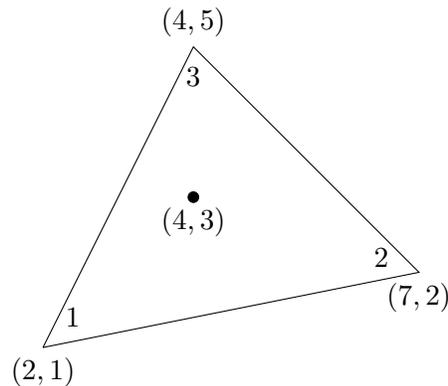


Figure 6.15 – Maillage et numéros de degrés de liberté

9. La figure 6.14 illustre un élément bidimensionnel de degré 3 de type Lagrange. Donner l'expression de la fonction  $\hat{\psi}_4(\xi, \eta)$ .
10. On donne le maillage de la figure 6.15. Le tableau de connectivité est :

<b>Connectivité</b>			
Tableau <i>connec</i>			
Élément	Noeud #1	Noeud #2	Noeud #3
1	1	4	5
2	1	5	2
3	2	5	6
4	2	6	3
5	4	7	8
6	4	8	5
7	5	8	9
8	5	9	6
9	7	10	11
10	7	11	8
11	8	11	12
12	8	12	9

Une condition de Dirichlet homogène a été imposée sur la paroi de droite (degrés de liberté 10 à 12) et des conditions de Neumann sur les 3 autres parois. On a utilisé des éléments triangulaires linéaires et le tableau d'adressage est le même que le tableau de connectivité.

Figure 6.16 – Élément  $K$ 

Après avoir résolu le problème, on a trouvé la solution :

$$U = \begin{bmatrix} 0,6362 \\ 0,7214 \\ 1,0000 \\ 0,5510 \\ 0,6248 \\ 0,8660 \\ 0,3181 \\ 0,3607 \\ 0,5000 \\ 0,0000 \\ 0,0000 \\ 0,0000 \end{bmatrix}$$

On considère le point  $\mathbf{P} = (2,25, 1,75)$  dans l'élément 12.

- a) Donnez l'expression de la transformation linéaire  $T^{K_{12}}$  qui envoie l'élément de référence sur le douzième élément.
  - b) Trouvez le point  $(\xi, \eta)$  de l'élément de référence qui est envoyé sur  $\mathbf{P}$ .
  - c) Évaluer  $u(\mathbf{P})$  et  $\frac{\partial u}{\partial x_1}(\mathbf{P})$ .
11. On a résolu un problème en éléments finis et on a trouvé une solution numérique  $u_h(x_1, x_2)$  telle que  $u_h(2, 1) = 5$ ,  $u_h(7, 2) = 53$  et  $u_h(4, 5) = 41$ . Évaluer  $u_h(4, 3)$  et  $\frac{\partial u_h}{\partial x_1}(4, 3)$  en vous servant de la figure 6.16.

**Remarque :** Il vous faudra inverser la transformation géométrique  $T^K$  appropriée pour déterminer à quelle coordonnée de l'élément de référence correspond le point  $(4, 3)$ .

## Chapitre 7

# Analyse de convergence

Nous nous proposons dans ce chapitre d'étudier comment convergent les méthodes d'éléments finis pour les problèmes elliptiques et à quel ordre. Nous compléterons le chapitre avec quelques exemples.

### 7.1 Bases théoriques

Nous avons vu au chapitre 3 les conditions d'existence et d'unicité d'une solution  $u$  d'un problème de la forme :

♠ trouver une fonction  $u \in V$  telle que :

$$a(u, w) = l(w) \quad \forall w \in V \quad (7.1)$$

Nous supposerons dans ce qui suit que les hypothèses du théorème de Lax-Milgram sont vérifiées à savoir que la forme linéaire  $l$  est continue et que la forme bilinéaire  $a$  est continue et coercive sur  $V$ .

Nous avons vu comment construire des solutions numériques par la méthode des éléments finis pour obtenir une approximation  $u_h$  de cette solution unique. Cela revient à construire un sous-espace de dimension finie  $V_h$  de  $V$  et à calculer  $u_h \in V_h$  solution de :

$$a(u_h, w_h) = l(w_h) \quad \forall w_h \in V_h \quad (7.2)$$

Le sous-espace  $V_h$  a pour base les fonctions de Ritz  $\phi_i(\mathbf{x})$  associées aux  $nndl$  degrés de liberté du maillage. On a en fait :

$$V_h = \{w_h | w_h = \sum_{i=1}^{nndl} \alpha_i \phi_i(\mathbf{x}), \alpha_i \in R\}$$

qui est un espace de dimension finie. Puisque l'on a construit des fonctions de Ritz  $\phi_i(\mathbf{x})$  appartenant à  $V$ , on a par construction  $V_h \subset V$ . Mentionnons toutefois que cette inclusion n'est pas absolument

nécessaire comme en font foi les éléments dits non conformes qui sont toutefois hors de notre propos pour le moment.

L'indice  $h$  réfère à la taille des éléments du maillage. Plus  $h$  est petit, plus les éléments sont petits et plus ils sont nombreux, et plus il y a de degrés de liberté. La dimension de  $V_h$  augmente donc lorsque  $h$  diminue. Pour un élément  $K$  du maillage, on note  $h^K$  la taille de l'élément. En dimension 1,  $h^K$  est simplement la longueur de l'élément mais en dimension 2 ou 3, on peut définir cette taille de plusieurs façons. En général, on pose :

$$h^K = \max_{\mathbf{x}, \mathbf{y} \in K} \|\mathbf{x} - \mathbf{y}\|_e$$

où  $\|\cdot\|_e$  est la norme euclidienne. C'est donc la distance maximale entre 2 points de l'élément. On pose ensuite :

$$h = \max_K h^K$$

Intuitivement, plus l'espace  $V_h$  est riche, plus la solution numérique  $u_h$  devrait se rapprocher de la solution  $u$ . Pour enrichir  $V_h$  on peut soit augmenter le nombre d'éléments, soit augmenter le degré des polynômes utilisés dans chaque élément. Il reste à s'assurer dans quelle(s) condition(s) on a bien convergence de  $u_h$  vers  $u$ , et s'il y a convergence, à quel ordre, etc. Le nombre d'éléments ainsi que le degré des polynômes utilisés auront des rôles importants à jouer. Pour fixer les idées, rappelons un peu de terminologie qui nous sera utile par la suite.

**Définition 7.1**

On dit que la solution numérique  $u_h$  converge vers  $u$  dans  $V$  si :

$$\lim_{h \rightarrow 0} \|u - u_h\|_V = 0$$

Lorsque  $h \rightarrow 0$ , la taille de tous les éléments tend vers 0. Il est important de remarquer que la norme utilisée est celle de l'espace  $V$ .

**Définition 7.2**

On dit que la solution numérique  $u_h$  converge à l'ordre  $p$  vers  $u$  dans  $V$  s'il existe une constante  $C$  telle que :

$$\|u - u_h\|_V \simeq Ch^p$$

lorsque  $h$  tend vers 0.

◆ **Remarque 7.1**

Dans ce qui suit, le même symbole  $C$  désignera diverses constantes indépendantes de  $h$ . ◆

★ **Théorème 7.1 (Lemme de Céa)**

Si  $a$  est une forme bilinéaire continue et coercive et si  $u$  et  $u_h$  dénotent les solutions respectives de 7.1 et 7.2, alors il existe une constante  $C$  indépendante de  $h$  telle que :

$$\|u - u_h\|_V \leq C \inf_{w_h \in V_h} \|u - w_h\|_V \quad (7.3)$$

**Démonstration :**

Ce résultat très simple mais aussi très important est dû à Céa [11]. On a immédiatement en vertu des problèmes 7.1 et 7.2 que :

$$a(u, w) = l(w) \quad \forall w \in V$$

$$a(u_h, w_h) = l(w_h) \quad \forall w_h \in V_h$$

Puisque  $V_h \subset V$ , on peut prendre  $w = w_h$  dans la première équation et soustraire pour obtenir :

$$a(u - u_h, w_h) = 0 \quad \forall w_h \in V_h \quad (7.4)$$

En particulier :

$$a(u - u_h, u_h) = 0$$

et on en conclut que :

$$a(u - u_h, u - u_h) = a(u - u_h, u - w_h) \quad \forall w_h \in V_h$$

La coercivité et la continuité de la forme bilinéaire  $a$  nous permettent alors d'écrire :

$$\alpha \|u - u_h\|_V^2 \leq C \|u - u_h\|_V \|u - w_h\|_V \quad \forall w_h \in V_h$$

ce qui entraîne que :

$$\|u - u_h\|_V \leq \frac{C}{\alpha} \|u - w_h\|_V \quad \forall w_h \in V_h$$

Puisque cette dernière inégalité est vraie quel que soit  $w_h \in V_h$ , on a le résultat. ★

L'interprétation de ce résultat est fondamentale. Nous venons en fait de ramener l'analyse de convergence de la solution éléments finis à un problème classique d'interpolation. En effet, la relation 7.3 affirme qu'à une constante près, l'erreur commise  $\|u - u_h\|_V$  est plus petite que la meilleure interpolation possible de  $u$  par des fonctions de  $V_h$ .

Notre intuition de départ se précise donc. En effet, plus l'espace  $V_h$  est riche, meilleure est la qualité de l'interpolation de  $u$  par des fonctions de  $V_h$  et en vertu de la relation 7.3, meilleure sera notre solution par éléments finis. La richesse de l'espace discret est bien sûr fonction du nombre d'éléments du maillage mais aussi du nombre de noeuds de calcul de chaque élément. Il est en effet clair que pour un même nombre d'éléments, une interpolation linéaire sera moins riche qu'une interpolation quadratique qui sera elle-même moins riche qu'une interpolation cubique, etc. La différence entre ces différentes approximations se manifesterà au niveau de l'ordre de convergence.

## 7.2 Quelques exemples

### ► Exemple 7.1

Supposons que le problème 7.1 soit suffisamment simple pour que sa solution  $u$  soit dans  $V_h$  (ce qui revient à supposer que l'espace discret  $V_h$  est suffisamment riche pour contenir la solution  $u$ ).

De l'équation 7.4, la solution numérique vérifie :

$$a(u - u_h, w_h) = 0 \quad \forall w_h \in V_h$$

En particulier, on peut prendre  $w_h = u - u_h$  ce qui est possible car  $u \in V_h$  et donc  $u - u_h \in V_h$ . On obtient ainsi :

$$a(u - u_h, u - u_h) = 0$$

La coercivité de la forme bilinéaire  $a$  entraîne alors que  $\|u - u_h\|_V = 0$  et donc que  $u = u_h$ .

Ce résultat trivial a quand même une grande importance pratique. Si un problème possède une solution telle que l'on peut l'interpoler exactement (sans erreur) dans  $V_h$ , alors la solution par éléments finis  $u_h$  sera aussi exacte. On se sert régulièrement de ce résultat pour vérifier un programme d'éléments finis. Au besoin, on construit de façon artificielle un problème dont la solution est dans l'espace discret  $V_h$  et on calcule par éléments finis cette solution. Tant que la solution analytique n'est pas reproduite exactement (à la précision machine), on conclut que le programme a une ou des erreur(s). Cette stratégie est fort utile pour « déboguer » un programme. ◀

### ★ Théorème 7.2

Si  $a$  est une forme bilinéaire symétrique, continue et coercive et si  $u$  et  $u_h$  dénotent les solutions respectives de 7.1 et 7.2, alors on a :

$$I(u) = \inf_{w \in V} I(w) \quad \text{et} \quad I(u_h) = \inf_{w_h \in V_h} I(w_h) \quad (7.5)$$

où  $I(w) = \frac{1}{2}a(w, w) - l(w)$ . De plus, on a :

$$I(u) \leq I(u_h) \quad (7.6)$$

★

Nous avons déjà démontré au chapitre 3 (voir la relation 3.5) l'équivalence entre les problèmes 7.1 et 7.2 et les problèmes de minimisation 7.5. La démonstration de l'inégalité 7.6 est immédiate puisque l'on a supposé que  $V_h \subset V$ . Ce qu'apporte en plus ce dernier résultat est que lorsqu'il y a convergence,  $I(u_h) \rightarrow I(u)$  en décroissant. Puisque  $I(u)$  est souvent liée à une énergie (fonctionnelle d'énergie), cela signifie que la solution par éléments finis surestimera l'énergie du système.

Terminons enfin avec le résultat le plus important qui porte sur l'ordre de convergence d'une solution éléments finis.

★ **Théorème 7.3 (Ordre de convergence)**

Soit  $u$  et  $u_h$  les solutions respectives des problèmes continu 7.1 et discret 7.2. On suppose de plus que la solution  $u$  du problème continu soit suffisamment régulière. Supposons enfin que l'espace  $V_h \subset H^m(\Omega)$  soit tel que sa restriction à chaque élément  $K$  contienne les polynômes de degré  $k$ , alors :

$$\|u - u_h\|_{m,\Omega} \leq Ch^{k+1-m} \|u\|_{k+1,\Omega} \quad (7.7)$$

**Preuve :** voir Ciarlet, réf. [9] ★

La méthode des éléments finis permet justement la construction de ce type d'espace  $V_h$  de fonctions polynômiales par élément. Ce dernier résultat nous donne alors l'ordre de convergence de la méthode d'éléments finis utilisée lorsqu'on choisit des polynômes de degré  $k$  sur chaque élément. Nous pouvons maintenant revenir sur certains des problèmes pour lesquels nous avons calculé une solution par éléments finis.

► **Exemple 7.2**

Pour les équations aux dérivées partielles d'ordre 2, l'espace de base est  $H^1(\Omega)$  et on doit utiliser la norme correspondante ( $m = 1$  dans l'inégalité 7.7). En utilisant une discrétisation par éléments finis linéaires ( $k = 1$ ), on a :

$$\|u - u_h\|_V = \|u - u_h\|_{1,\Omega} = (\|u - u_h\|_{0,\Omega}^2 + \|\nabla u - \nabla u_h\|_{0,\Omega}^2)^{1/2} \leq Ch \|u\|_{2,\Omega}$$

On constate donc une convergence linéaire en norme  $H^1(\Omega)$  en supposant que la solution  $u$  soit dans  $H^2(\Omega)$  ce qui n'est pas toujours le cas. On sait en effet que la solution  $u$  est dans  $H^1(\Omega)$  et que  $H^2(\Omega) \subset H^1(\Omega)$ . Rien ne nous assure donc *a priori* que  $u$  soit dans  $H^2(\Omega)$ . C'est ce que nous signifions lorsque nous supposons que  $u$  est suffisamment régulière dans l'énoncé du théorème.

Notons que le résultat est également vrai en norme  $L^2(\Omega)$  et que :

$$\|u - u_h\|_{0,\Omega} \leq Ch^2 \|u\|_{2,\Omega}$$

et que la convergence est alors quadratique dans cette norme. Bien que valide, ce dernier résultat ne donne pas une idée exacte de l'ordre de convergence de la solution puisque nous utilisons le mauvais instrument de mesure, c.-à-d. la mauvaise norme. On peut toutefois, puisque la différence entre la norme  $H^1(\Omega)$  et la norme  $L^2(\Omega)$  n'est rien d'autre que la norme  $L^2(\Omega)$  des dérivées partielles  $\|\nabla u - \nabla u_h\|_{0,\Omega}$ , conclure que ce sont ces dérivées partielles qui convergent moins vite. On a en fait :

$$\|\nabla u - \nabla u_h\|_{0,\Omega} \leq Ch \|u\|_{2,\Omega}$$

Nous avons déjà remarqué lors de certains exemples numériques que les dérivées (ordinaires ou partielles) de  $u$  étaient plus difficile à approcher. Ce résultat le confirme de manière théorique.

Si on passe maintenant à des éléments quadratiques ( $k = 2$ ), on trouve une convergence quadratique :

$$\|u - u_h\|_V = \|u - u_h\|_{1,\Omega} \leq C h^{2+1-1} \|u\|_{3,\Omega}$$

ce qui exige cependant encore plus de régularité sur  $u$  ( $u \in H^3(\Omega)$ ). ◀

► **Exemple 7.3**

Pour les problèmes d'ordre 4, nous n'avons vu qu'un seul élément de degré 3 ( $k = 3$ ) en dimension 1. Cela nous donne en norme  $H^2(\Omega)$  ( $m = 2$ ) :

$$\|u - u_h\|_V = \|u - u_h\|_{2,\Omega} \leq C h^{3+1-2} \|u\|_{4,\Omega} = C h^2 \|u\|_{4,\Omega}$$

c.-à-d. une convergence d'ordre 2 (quadratique). Tout comme précédemment, on a également :

$$\|u - u_h\|_{2,\Omega} \leq C h^2 \|u\|_{4,\Omega}$$

$$\|u - u_h\|_{1,\Omega} \leq C h^3 \|u\|_{4,\Omega}$$

$$\|u - u_h\|_{0,\Omega} \leq C h^4 \|u\|_{4,\Omega}$$

et ici encore, on conclut que pour les problèmes d'ordre 4, les dérivées secondes convergent en norme  $L^2(\Omega)$  moins vite (à l'ordre 2) que les dérivées premières qui elles-mêmes convergent moins vite (à l'ordre 3) que  $\|u - u_h\|_{0,\Omega}$  qui converge à l'ordre 4. ◀

► **Exemple 7.4**

Illustrons maintenant concrètement comment on peut vérifier ces ordres de convergence théoriques. Cela est souvent utile pour s'assurer qu'un programme d'éléments finis a le bon comportement et donc pas d'erreur de programmation.

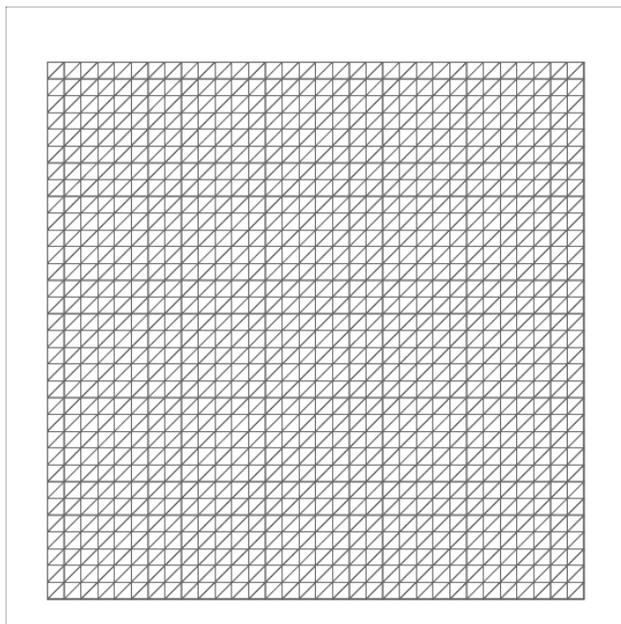
Pour ce faire, on se donne une solution analytique dont l'expression est suffisamment complexe pour ne pas appartenir à l'espace de discrétisation  $V_h$ . Dans cet exemple : on pose :

$$u(x) = x_1 x_2 (1 - x_1)(1 - x_2) \arctan \left( 20 \frac{(x_1 + x_2)}{\sqrt{2}} - 16 \right) \quad (7.8)$$

et on va résoudre l'équation de Laplace :

$$-\nabla^2 u(x) = f(x)$$

sur le carré  $[0, 1]^2$  et en choisissant comme terme de droite  $f(x)$  précisément le laplacien de la fonction 7.8 de sorte que celle-ci est bien la solution de notre problème. C'est ce que l'on appelle une *solution manufacturée* au sens de Roache [33]. Le choix d'une telle fonction est *a priori* arbitraire mais il est recommandé d'essayer de choisir des fonctions qui représentent au mieux les comportements observés sur de vrais problèmes.

Figure 7.1 – Maillage de  $32 \times 32$  éléments

On vérifie facilement que les conditions aux limites sont nulles sur toute la frontière du domaine. Notre stratégie consiste à obtenir des solutions à ce problème par la méthode des éléments finis sur des maillages de plus en plus fins. Nous avons choisi des maillages réguliers de 4 par 4, 8 par 8, 16 par 16, 32 par 32, 64 par 64 et enfin de 128 par 128 éléments. Le maillage de 32 par 32 éléments est illustré à la figure 7.1. Pour chaque maillage et chaque solution numérique  $u_h$ , on calcule les normes  $\|u - u_h\|_{1,\Omega}$  et  $\|u - u_h\|_{0,\Omega}$  de l'erreur. Rappelons que l'on aurait pu utiliser pour ce problème la norme  $|u - u_h|_{1,\Omega}$  (au lieu de la norme  $\|u - u_h\|_{1,\Omega}$ ) car on travaille en fait dans  $H_0^1(\Omega)$ . De plus, on calcule la norme  $\|u - u_h\|_{0,\Omega}$  seulement pour illustrer plus en détails le comportement de l'erreur. Enfin, nous avons considéré des approximations linéaires et quadratiques pour souligner l'effet du degré des polynômes sur la convergence.

### Approximation linéaire

Dans ce premier cas, on choisit des approximations linéaires (voir la figure 6.2) sur chaque élément. Nous notons  $u_h^1$  la solution numérique obtenue. On sait alors que :

$$\|u - u_h^1\|_{1,\Omega} \simeq Ch$$

ou encore, en prenant le logarithme naturel de chaque côté :

$$\ln(\|u - u_h^1\|_{1,\Omega}) \simeq 1 \ln h + \ln C$$

Sur échelle logarithmique, on obtient ainsi une droite de pente 1 (qui est l'ordre de convergence). Ce résultat est bien sûr asymptotique en ce sens qu'il n'est vrai que pour des valeurs de  $h$  suffisamment

petites. C'est ce que nous constatons dans cet exemple à la figure 7.2 tracée à partir des données de la table suivante :

Convergence : cas linéaire			
Maillage	$h$	$\ u - u_h^1\ _{1,\Omega}$	$\ u - u_h^1\ _{0,\Omega}$
$4 \times 4$	0,25	0,464 812	$0,1795 \times 10^{-1}$
$8 \times 8$	0,125	0,323 407	$0,8231 \times 10^{-2}$
$16 \times 16$	0,0625	0,200 552	$0,2636 \times 10^{-2}$
$32 \times 32$	0,03125	0,110 145	$0,7640 \times 10^{-3}$
$64 \times 64$	0,015625	0,056 758	$0,2025 \times 10^{-3}$
$128 \times 128$	0,0078125	0,028 613	$0,5146 \times 10^{-4}$

On remarque que pour les grandes valeurs de  $h$ , la courbe de  $\|u - u_h^1\|_{1,\Omega}$  a une pente différente de 1 mais au fur et à mesure que la valeur de  $h$  diminue, la pente devient très voisine de 1. On peut vérifier aisément que la pente est bien 1 en se servant de la table précédente.

Si on travaille en norme  $\|u - u_h\|_{0,\Omega}$ , on remarque que conformément à la relation 7.7, on a une courbe de pente 2. Ceci illustre une fois de plus que les dérivées de la fonctions  $u(x)$  convergent moins vite que la fonction  $u(x)$  elle-même.

### Approximation quadratique

Dans ce second cas, on choisit des approximations quadratiques sur chaque élément (voir la figure 6.2). Nous notons  $u_h^2$  la solution numérique obtenue. On sait alors que :

$$\|u - u_h^2\|_{1,\Omega} \simeq Ch^2$$

et on obtient cette fois une droite de pente 2 sur une échelle logarithmique :

$$\ln(\|u - u_h^2\|_{1,\Omega}) \simeq 2 \ln h + \ln C$$

Le tableau suivant résume les résultats obtenus :

Convergence : cas quadratique			
Maillage	$h$	$\ u - u_h^2\ _{1,\Omega}$	$\ u - u_h^2\ _{0,\Omega}$
$4 \times 4$	0,25	0,242 390	$0,5179 \times 10^{-2}$
$8 \times 8$	0,125	0,146 069	$0,3043 \times 10^{-2}$
$16 \times 16$	0,0625	0,052 467	$0,3694 \times 10^{-3}$
$32 \times 32$	0,03125	0,015 405	$0,3406 \times 10^{-4}$
$64 \times 64$	0,015625	0,004 136	$0,4345 \times 10^{-5}$
$128 \times 128$	0,0078125	0,001 056	$0,5477 \times 10^{-6}$

On observe le même comportement que dans le cas linéaire si ce n'est que les ordres de convergence sont maintenant 2 pour la norme  $\|u - u_h^2\|_{1,\Omega}$  et 3 pour la norme  $\|u - u_h^2\|_{0,\Omega}$  (voir la figure 7.2). On remarque de plus que, pour un maillage donné, les erreurs sont beaucoup plus faibles si on utilise une approximation quadratique. En fait, la solution obtenue sur un maillage avec une approximation

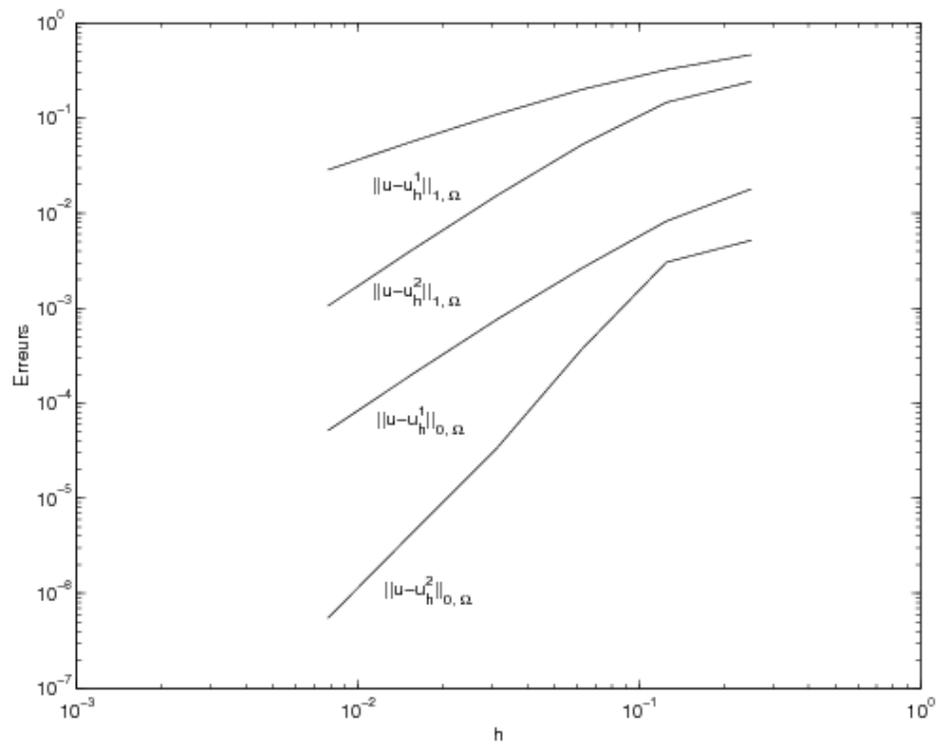


Figure 7.2 – Erreurs commises pour les approximations linéaire et quadratique

quadratique est toujours plus précise que celle obtenue avec une approximation linéaire sur le maillage ayant 4 fois plus d'éléments. ◀

► **Exemple 7.5**

Pour illustrer de manière plus pragmatique comment on peut vérifier la convergence d'une solution numérique, nous reprenons le problème thermique du chapitre 6 page 160. On retrouvera la géométrie et les conditions aux limites à la figure 6.12.

On peut s'assurer d'une plus grande précision en résolvant le même problème sur une série de maillages de plus en plus fins. Chaque maillage comporte 4 fois plus d'éléments que le précédent et est obtenu en divisant chaque arête du maillage en 2, multipliant ainsi le nombre d'éléments par 4. Les maillages utilisés sont présentés à la figure 7.3 et comportent respectivement 312, 1248, 4992 et 19968 éléments linéaires. On présente les isothermes correspondantes à la figure 7.4. On y constate une très légère évolution des isothermes même si le portrait global reste sensiblement le même. On peut davantage saisir ces variations à la figure 7.5 où nous avons effectué une coupe de la température sur l'axe  $y = 1$  correspondant au milieu de la partie inférieure de la géométrie. Nous avons également résolu le problème sur le maillage de 19968 éléments avec une approximation quadratique et cette solution nous servira de « solution exacte » (ce qui n'est pas vraiment le cas). On constate quand même que l'erreur est divisée par un facteur 2 sur chaque maillage puisque la valeur de  $h$  est divisée par 2 et que nous avons utilisé un élément linéaire. ◀

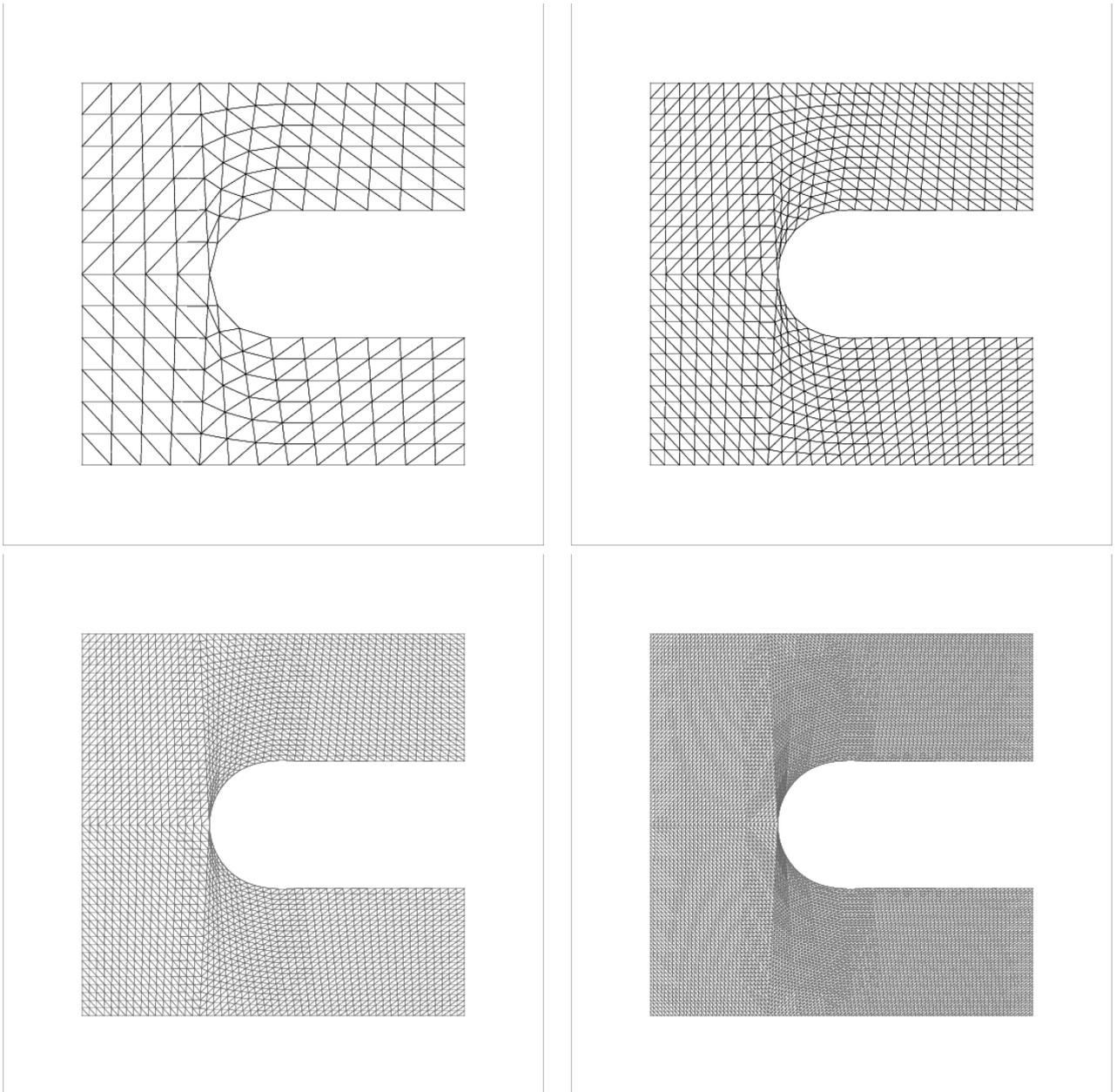


Figure 7.3 – Transfert thermique : maillages utilisés

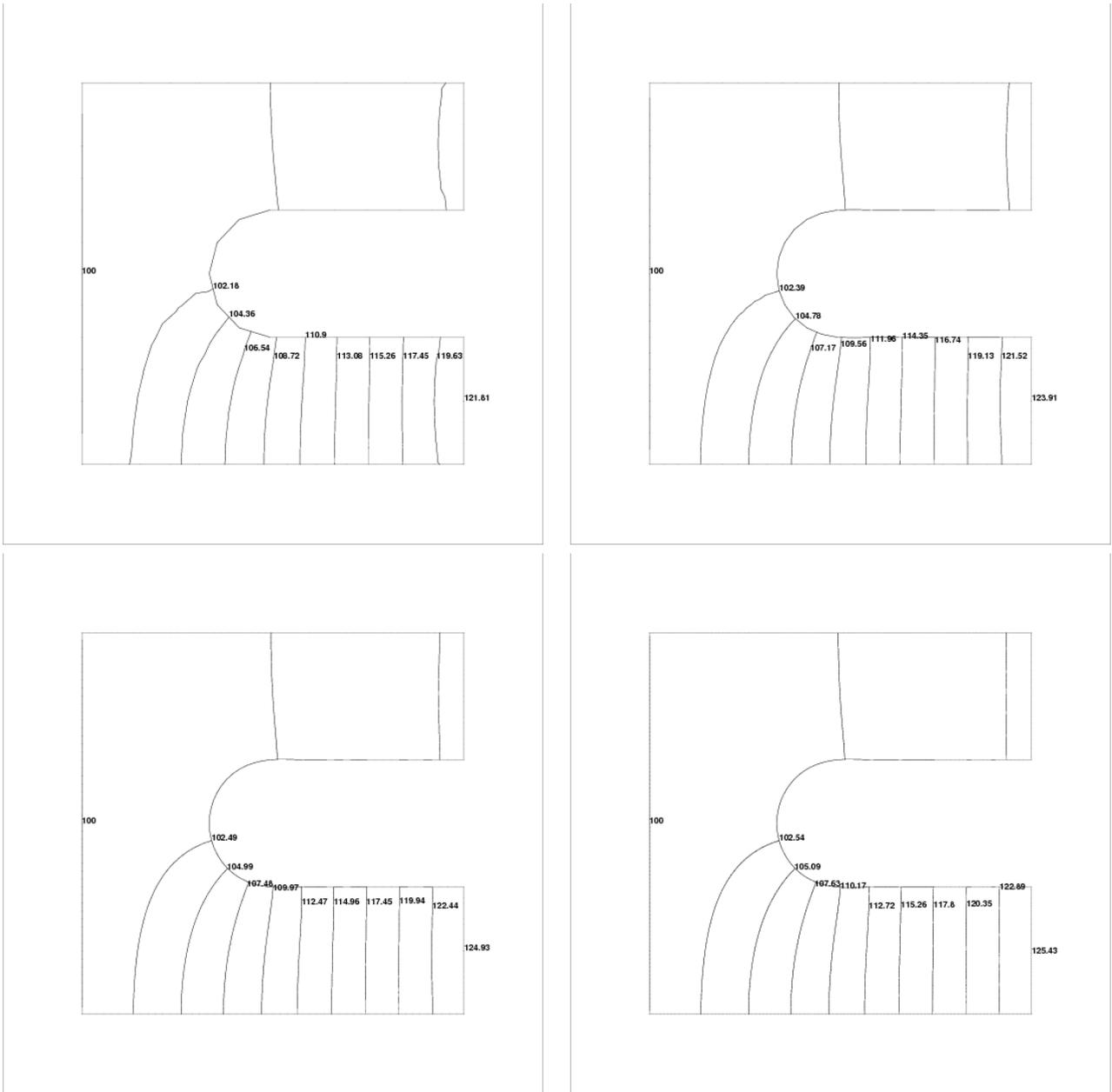


Figure 7.4 – Transfert thermique : isothermes

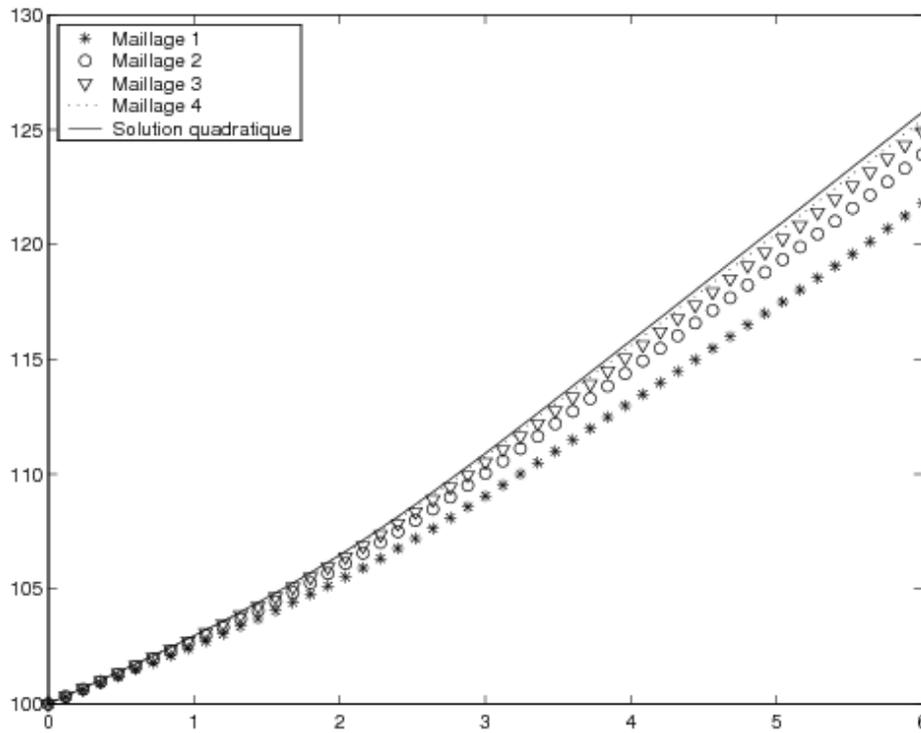


Figure 7.5 – Coupe de la température sur l'axe  $y = 1$



## Chapitre 8

# Problèmes non linéaires

Nous n'avons jusqu'à maintenant considéré que les problèmes linéaires qui sont certes importants mais non suffisants pour les applications. En effet, la nature, ou plus précisément les modèles que nous utilisons pour décrire les phénomènes naturels, est le plus souvent non linéaire. La théorie développée dans les premiers chapitres n'est malheureusement valide que dans le cas linéaire. En effet, le théorème de Lax-Milgram permet de montrer l'existence et l'unicité de problèmes de la forme :

♠ trouver une fonction  $u \in V$  telle que :

$$a(u, w) = l(w) \quad \forall w \in V \quad (8.1)$$

où  $l$  et  $a$  sont respectivement des formes linéaire et bilinéaire. Ce ne sera plus le cas dans ce chapitre.

Bien que la théorie des problèmes non linéaires soit beaucoup plus complexe, nous nous limiterons à une généralisation formelle de la théorie développée dans le cas linéaire. Il faudra toutefois être conscients que les espaces fonctionnels impliqués sont parfois plus complexes que les espaces  $H^1(\Omega)$  et  $H^2(\Omega)$  rencontrés jusqu'à maintenant.

Le plus simple est de considérer immédiatement un exemple d'équation aux dérivées partielles non linéaire :

$$p(\mathbf{x}, u(\mathbf{x}))u(\mathbf{x}) - \nabla \cdot (q(\mathbf{x}, u(\mathbf{x}))\nabla u(\mathbf{x})) = r(\mathbf{x}, u(\mathbf{x}))$$

auquel on ajoute par exemple des conditions aux limites essentielles sur  $u$ . On remarque la dépendance des fonctions  $p$ ,  $q$  et  $r$  sur la variable inconnue  $u$ . C'est ce qui rend le problème non linéaire. La non-linéarité pourra aussi, comme nous le verrons, provenir des conditions aux limites.

Nous procéderons comme dans le cas linéaire. Multiplions par une fonction test  $w(\mathbf{x})$  dans un espace fonctionnel  $V$  et intégrons par parties sur le domaine  $\Omega$ . On obtient ainsi :

$$\int_{\Omega} p(\mathbf{x}, u(\mathbf{x}))u(\mathbf{x}) w(\mathbf{x}) + q(\mathbf{x}, u(\mathbf{x}))\nabla u(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, dv - \int_{\Gamma} (q(\mathbf{x}, u(\mathbf{x}))\nabla u(\mathbf{x})) \cdot \mathbf{n} w \, ds = \int_{\Omega} r(\mathbf{x}, u(\mathbf{x})) w(\mathbf{x}) \, dv$$

Puisque nous avons supposé des conditions aux limites essentielles sur toute la frontière, l'intégrale de bord s'annule et il reste :

♠ trouver une fonction  $u(\mathbf{x}) \in V$  telle que :

$$\int_{\Omega} p(\mathbf{x}, u(\mathbf{x}))u(\mathbf{x}) w(\mathbf{x}) + q(\mathbf{x}, u(\mathbf{x}))\nabla u(\mathbf{x}) \cdot \nabla w(\mathbf{x}) dv = \int_{\Omega} r(\mathbf{x}, u(\mathbf{x}))w(\mathbf{x}) dv \quad \forall w(\mathbf{x}) \in V$$

On constate immédiatement que le terme de gauche de cette expression ne représente pas une forme bilinéaire en raison des coefficients  $p(\mathbf{x}, u)$  et  $q(\mathbf{x}, u)$ . Par conséquent, le théorème de Lax-Milgram ne peut donc pas s'appliquer ici.

Pour résoudre, on devra linéariser le problème c'est-à-dire en ramener la solution à un problème linéaire ou plus précisément à une suite de problèmes linéaires. La convergence de cette suite de problèmes linéaires vers la solution du problème non linéaire de départ n'est toutefois nullement garantie. On voit immédiatement l'analogie avec les systèmes de fonctions algébriques non linéaires :

$$\mathbf{f}(\mathbf{x}) = 0$$

que l'on peut résoudre (voir Fortin, réf. [17]) par des méthodes de points fixes, par la méthode de Newton, etc. Toutes ces méthodes se généralisent assez facilement aux problèmes variationnels.

## 8.1 Rappel sur les systèmes d'équations non linéaires

Dans cette section, nous rappelons quelques techniques de base pour la résolution des équations non linéaires. Le problème consiste à trouver le ou les vecteurs  $\mathbf{x} = (x_1, x_2, x_3, \dots, x_n)$  vérifiant les  $n$  équations non linéaires suivantes :

$$\mathbf{f}(\mathbf{x}) = 0$$

ou plus explicitement :

$$\begin{cases} f_1(x_1, x_2, x_3, \dots, x_n) = 0 \\ f_2(x_1, x_2, x_3, \dots, x_n) = 0 \\ f_3(x_1, x_2, x_3, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, x_3, \dots, x_n) = 0 \end{cases} \quad (8.2)$$

où les  $f_i$  sont des fonctions de  $n$  variables que nous supposons différentiables. Contrairement aux systèmes linéaires, il n'y a pas de condition simple associée aux systèmes non linéaires qui permette d'assurer l'existence et l'unicité de la solution.

Les méthodes de résolution des systèmes non linéaires sont nombreuses. Nous présentons ici un cadre relativement général pour les méthodes itératives avec comme cas particulier important la *méthode de Newton*. La présentation inclura toutefois des variantes de cette méthode de même qu'un aperçu des méthodes de points fixes.

L'application de la méthode de Newton à un système de deux équations non linéaires est suffisante pour illustrer le cas général. Considérons donc le système :

$$\begin{cases} f_1(x_1, x_2) = 0 \\ f_2(x_1, x_2) = 0 \end{cases}$$

Soit  $(x_1^0, x_2^0)$ , une approximation initiale de la solution de ce système. Cette approximation initiale est cruciale et doit toujours être choisie avec soin. Le but de ce qui suit est de déterminer une correction  $(\delta_{x_1}, \delta_{x_2})$  à  $(x_1^0, x_2^0)$  de telle sorte que :

$$\begin{aligned} f_1(x_1^0 + \delta_{x_1}, x_2^0 + \delta_{x_2}) &= 0 \\ f_2(x_1^0 + \delta_{x_1}, x_2^0 + \delta_{x_2}) &= 0 \end{aligned}$$

Pour déterminer  $(\delta_{x_1}, \delta_{x_2})$ , il suffit maintenant de faire un développement de Taylor en deux variables pour chacune des deux fonctions :

$$\begin{aligned} 0 &= f_1(x_1^0, x_2^0) + \frac{\partial f_1}{\partial x_1}(x_1^0, x_2^0) \delta_{x_1} + \frac{\partial f_1}{\partial x_2}(x_1^0, x_2^0) \delta_{x_2} + \dots \\ 0 &= f_2(x_1^0, x_2^0) + \frac{\partial f_2}{\partial x_1}(x_1^0, x_2^0) \delta_{x_1} + \frac{\partial f_2}{\partial x_2}(x_1^0, x_2^0) \delta_{x_2} + \dots \end{aligned}$$

Dans les relations précédentes, les pointillés désignent des termes d'ordre supérieur ou égal à deux et faisant intervenir les dérivées partielles d'ordre correspondant. Pour déterminer  $(\delta_{x_1}, \delta_{x_2})$ , il suffit de négliger les termes d'ordre supérieur et d'écrire :

$$\begin{aligned} \frac{\partial f_1}{\partial x_1}(x_1^0, x_2^0) \delta_{x_1} + \frac{\partial f_1}{\partial x_2}(x_1^0, x_2^0) \delta_{x_2} &= -f_1(x_1^0, x_2^0) \\ \frac{\partial f_2}{\partial x_1}(x_1^0, x_2^0) \delta_{x_1} + \frac{\partial f_2}{\partial x_2}(x_1^0, x_2^0) \delta_{x_2} &= -f_2(x_1^0, x_2^0) \end{aligned}$$

ou encore sous forme matricielle :

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_1^0, x_2^0) & \frac{\partial f_1}{\partial x_2}(x_1^0, x_2^0) \\ \frac{\partial f_2}{\partial x_1}(x_1^0, x_2^0) & \frac{\partial f_2}{\partial x_2}(x_1^0, x_2^0) \end{bmatrix} \begin{bmatrix} \delta_{x_1} \\ \delta_{x_2} \end{bmatrix} = - \begin{bmatrix} f_1(x_1^0, x_2^0) \\ f_2(x_1^0, x_2^0) \end{bmatrix}$$

Ce système linéaire s'écrit également sous une forme plus compacte :

$$J_f(x_1^0, x_2^0) \boldsymbol{\delta}_x = -\mathbf{f}(x_1^0, x_2^0) \tag{8.3}$$

où  $J_f(x_1^0, x_2^0)$  désigne la matrice des dérivées partielles ou *matrice jacobienne* évaluée au point  $(x_1^0, x_2^0)$ ,  $\boldsymbol{\delta}_x = (\delta_{x_1}, \delta_{x_2})$  est le vecteur des corrections relatives à chaque variable et où  $-\mathbf{f}(x_1^0, x_2^0)$  est le *vecteur résidu* évalué en  $(x_1^0, x_2^0)$ . Le déterminant de la matrice jacobienne est appelé le *jacobien*. Le jacobien doit bien entendu être différent de 0 pour que la matrice jacobienne soit inversible. On pose ensuite :

$$\begin{aligned} x_1^1 &= x_1^0 + \delta_{x_1} \\ x_2^1 &= x_2^0 + \delta_{x_2} \end{aligned}$$

qui est la nouvelle approximation de la solution du système non linéaire. On cherchera par la suite à corriger  $(x_1^1, x_2^1)$  d'une nouvelle quantité  $(\delta_x)$ , et ce jusqu'à la convergence.

De manière plus générale, on pose :

$$J_f(\mathbf{x}^k) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}^k) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}^k) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}^k) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}^k) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}^k) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}^k) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}^k) & \frac{\partial f_n}{\partial x_2}(\mathbf{x}^k) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}^k) \end{bmatrix}$$

c'est-à-dire la matrice jacobienne évaluée au point  $\mathbf{x}^k = (x_1^k, x_2^k, \dots, x_n^k)$ . De plus on pose :

$$\mathbf{f}(\mathbf{x}^k) = \begin{bmatrix} f_1(\mathbf{x}^k) \\ f_2(\mathbf{x}^k) \\ \vdots \\ f_n(\mathbf{x}^k) \end{bmatrix} \quad \delta_x = \begin{bmatrix} \delta_{x_1} \\ \delta_{x_2} \\ \vdots \\ \delta_{x_n} \end{bmatrix}$$

pour en arriver à l'algorithme général suivant extrait de la référence [17].

**Algorithme 8.1 :** *Méthode de Newton*

1. Étant donné  $\epsilon$ , un critère d'arrêt
2. Étant donné  $N$ , le nombre maximal d'itérations
3. Étant donné  $\mathbf{x}^0 = [x_1^0 \ x_2^0 \ \cdots \ x_n^0]^T$ , une approximation initiale de la solution du système
4. Résoudre le système linéaire :

$$J_f(\mathbf{x}^k) \delta_x = -\mathbf{f}(\mathbf{x}^k) \tag{8.4}$$

et poser :

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \delta_x$$

5. Si  $\frac{\|\delta_x\|}{\|\mathbf{x}^{k+1}\|} < \epsilon$  et  $\|\mathbf{f}(\mathbf{x}^{k+1})\| \leq \epsilon$  :
  - convergence atteinte
  - écrire la solution  $\mathbf{x}^{k+1}$
  - arrêt
6. Si le nombre maximal d'itérations  $N$  est atteint :
  - convergence non atteinte en  $N$  itérations
  - arrêt
7. Retour à l'étape 4



La méthode de Newton n'est qu'un cas particulier de schéma itératif pour la résolution du système 8.2. De façon plus générale, on peut écrire un schéma itératif sous la forme :

$$A_k(\mathbf{x}^{k+1} - \mathbf{x}^k) + \mathbf{f}(\mathbf{x}^k) = 0 \tag{8.5}$$

où  $A_k$  est une matrice non singulière. L'équation 8.5 a alors comme unique solution :

$$\mathbf{x}^{k+1} = \mathbf{x}^k - A_k^{-1} \mathbf{f}(\mathbf{x}^k) \tag{8.6}$$

La méthode de Newton n'est qu'un cas particulier où  $A_k = J_f(\mathbf{x}^k)$ .

◆ **Remarque 8.1**

Il est bien entendu que l'expression 8.6 ne signifie nullement qu'il faille inverser la matrice  $A_k$  ce qui serait terriblement coûteux. On résoudra en fait le système linéaire 8.5. ◆

Les variantes de la relation 8.5 sont nombreuses. L'idée de base étant toujours de diminuer les coûts de calcul et de mémoire. En effet, le calcul de la matrice jacobienne est une opération coûteuse qui doit être effectuée à chaque itération. On a alors imaginé des variantes qui contournent cette difficulté au moins en partie. Ces variantes dépendent de la forme de la fonction  $\mathbf{f}(\mathbf{x})$ . Par exemple, si on a :

$$\mathbf{f}(\mathbf{x}) = B(\mathbf{x})\mathbf{x} - \mathbf{g}(\mathbf{x}) \tag{8.7}$$

où  $B(\mathbf{x})$  est une matrice dont les coefficients peuvent dépendre eux-mêmes de la solution (mais pas forcément) et que l'on suppose inversible et  $\mathbf{g}(\mathbf{x})$  est une autre fonction non linéaire. En prenant  $A_k = B(\mathbf{x}^k)$ , l'équation 8.6 devient :

$$\mathbf{x}^{k+1} = \mathbf{x}^k - B^{-1}(\mathbf{x}^k) \left( B(\mathbf{x}^k)\mathbf{x}^k - \mathbf{g}(\mathbf{x}^k) \right) = B^{-1}(\mathbf{x}^k)\mathbf{g}(\mathbf{x}^k)$$

qui revient au système linéaire :

$$B(\mathbf{x}^k)\mathbf{x}^{k+1} = \mathbf{g}(\mathbf{x}^k) \tag{8.8}$$

En particulier, si  $B = I$ , on a la méthode des points fixes classique. Si la matrice  $B$  ne dépend pas de  $\mathbf{x}$ , la convergence de l'algorithme 8.8 vers la solution (le point fixe)  $\mathbf{r}$  est liée au rayon spectral  $\rho$  (voir Fortin [17]) de la matrice  $B^{-1}(J_g(\mathbf{r}))$  ( $J_g$  est la matrice jacobienne associée à la fonction  $g$ ). Plus précisément, on doit avoir :

$$\rho(B^{-1}(J_g(\mathbf{r}))) < 1$$

pour que l'algorithme puisse converger. Encore faut-il que l'estimation de départ  $\mathbf{x}^0$  ne soit pas trop loin de la solution. En pratique, il est extrêmement difficile de s'assurer que cette condition est satisfaite.

De nombreuses autres variantes entrent dans ce cadre général. Par exemple, on peut prendre  $A_k = J_f(\mathbf{x}^0) = A_0$  ce qui revient à fixer la matrice itérative  $A_k$  à la première matrice

jacobienne. Cela évite de réassembler et refactoriser une nouvelle matrice jacobienne à chaque nouvelle itération. Il peut en résulter un gain important en temps de calcul au risque d'une plus grande vulnérabilité au niveau des propriétés de convergence. On peut aussi remettre à jour cette matrice de temps à autre. Cette variante est parfois intéressante pour les problèmes instationnaires où il n'est pas forcément nécessaire de calculer la matrice jacobienne à chaque itération de chaque pas de temps.

## 8.2 Dérivée d'une fonctionnelle

Pour éventuellement appliquer la méthode de Newton à la résolution dans le cas d'une formulation variationnelle non linéaire, nous devons étendre le concept de dérivée aux fonctionnelles. C'est l'objectif des définitions suivantes.

### Définition 8.1

Soit  $G(v)$  une fonctionnelle définie sur un espace de Hilbert  $V$ . On dit que  $G$  est dérivable en  $v_0$  s'il existe une fonctionnelle linéaire de  $V$  dans  $\mathcal{R}$  notée  $D_v G(v_0)$  telle que :

$$G(v_0 + \delta_v) = G(v_0) + D_v G(v_0)[\delta_v] + \|\delta_v\| o(\delta_v) \quad (8.9)$$

où  $o(\delta_v)$  est une fonctionnelle vérifiant :

$$\lim_{\delta_v \rightarrow 0} o(\delta_v) = 0$$

Cette définition revient à supposer l'existence d'une forme de développement de Taylor puisque le dernier terme de droite doit être petit pour des  $\delta_v$  petits. Il reste à définir le sens de l'expression :

$$D_v G(v_0)[\delta_v]$$

### Définition 8.2

Soit  $G(v)$  une fonctionnelle dérivable en  $v_0$ . La dérivée de Gateaux de la fonctionnelle  $G$  par rapport à  $v$  en  $v_0$  et dans la direction  $\delta_v$  est définie par :

$$D_v G(v_0)[\delta_v] = \frac{\partial G}{\partial v}(v_0)[\delta_v] = \lim_{\epsilon \rightarrow 0} \frac{G(v_0 + \epsilon \delta_v) - G(v_0)}{\epsilon} = \frac{d}{d\epsilon} [G(v_0 + \epsilon \delta_v)] \Big|_{\epsilon=0} \quad (8.10)$$

si cette limite existe. Si la limite existe  $\forall \delta_v \in V$ , alors on dit que  $G$  possède une dérivée de Gateau en  $v_0$ .

On reconnaît là une dérivée directionnelle. On nomme parfois cette dérivée la *première variation de G*. Le terme de droite contient une dérivée classique par rapport à une variable réelle  $\epsilon$  et est plus simple à calculer.<sup>1</sup>

◆ **Remarque 8.2**

Propriétés de la dérivée de Gateaux :

En vue des applications des prochains chapitres, nous nous placerons dans un cadre légèrement plus général. Nous considérerons en effet des fonctionnelles allant d'un espace vectoriel normé  $V$  à un autre espace vectoriel normé  $W$ . La dérivée de Gateaux est alors une fonctionnelle linéaire de  $V$  dans  $W$ . Dans beaucoup de situations,  $W$  sera tout simplement l'ensemble des réels  $R$ .

1. **Dérivée d'une somme** : Si  $G_1$  et  $G_2$  sont des applications de  $V$  dans  $W$  et si  $G(v) = G_1(v) + G_2(v)$ , alors :

$$D_v(G_1 + G_2)(v_0)[\delta_v] = D_v G_1(v_0)[\delta_v] + D_v G_2(v_0)[\delta_v] \tag{8.11}$$

2. **Dérivée d'un produit** : Si  $G_1$  et  $G_2$  sont des applications de  $V$  dans  $W$  et si on pose  $G(v) = G_1(v) \cdot G_2(v)$ , alors :

$$D_v(G_1 \cdot G_2)(v_0)[\delta_v] = D_v G_1(v_0)[\delta_v] \cdot G_2(v_0) + G_1(v_0) \cdot D_v G_2(v_0)[\delta_v] \tag{8.12}$$

On notera au passage que la notation  $(\cdot)$  peut désigner toute forme de produit définie sur  $W$ .

3. **Dérivée de la composition de deux fonctionnelles** : Si  $G_1$  est une fonctionnelle de  $V$  dans  $U$  (un espace vectoriel) et si  $G_2$  est une fonctionnelle de  $U$  dans  $W$ , alors on peut composer  $G_1$  et  $G_2$  et ainsi définir une fonctionnelle  $G$  de  $V$  dans  $W$  par la composition suivante :  $v \rightarrow G_1(v) = u \in U \rightarrow G_2(u) = G_2(G_1(v)) = G(v) \in W$ . La dérivée de Gateaux est alors une fonctionnelle linéaire de  $V$  dans  $W$  définie par :

$$D_v(G_2(G_1(v_0)))[\delta_v] = D_u G_2(u_0) [(D_v G_1(v_0)[\delta_v]] \quad \text{où } u_0 = G_1(v_0) \tag{8.13}$$



### 8.3 Application aux formulations variationnelles

Considérons le problème non linéaire :

$$\begin{aligned} (u(\mathbf{x}))^\alpha - \nabla \cdot (q(\mathbf{x}) \nabla u(\mathbf{x})) &= r(\mathbf{x}) \quad \text{dans } \Omega \\ u(\mathbf{x}) &= 0 \quad \text{sur } \Gamma \end{aligned} \tag{8.14}$$

---

1. René Eugène Gateaux (1889-1914) était un jeune mathématicien français tué au combat au début de la première guerre mondiale et dont les travaux ont été repris et généralisés par Paul Pierre Lévy et publiés à partir de 1919. N.B. D'après son acte de naissance, son nom s'orthographe sans accent circonflexe.

Pour simplifier l'exposé, nous supposons des conditions essentielles homogènes sur  $u$ , le cas non homogène ne posant aucune difficulté supplémentaire. La non-linéarité provient du premier terme de gauche affublé d'un exposant  $\alpha$  pour le moment quelconque (mais différent de 0 ou 1). Une formulation variationnelle pour ce problème est alors :

♠ Trouver  $u(\mathbf{x}) \in H_0^1(\Omega)$  telle que :

$$\int_{\Omega} ((u(\mathbf{x}))^\alpha w(\mathbf{x}) + q(\mathbf{x}) \nabla u(\mathbf{x}) \cdot \nabla w(\mathbf{x})) \, dv = \int_{\Omega} r(\mathbf{x}) w(\mathbf{x}) \, dv \quad \forall w(\mathbf{x}) \in H_0^1(\Omega)$$

On pose alors :

$$R(u, w) = \int_{\Omega} ((u(\mathbf{x}))^\alpha w(\mathbf{x}) + q(\mathbf{x}) \nabla u(\mathbf{x}) \cdot \nabla w(\mathbf{x})) \, dv - \int_{\Omega} r(\mathbf{x}) w(\mathbf{x}) \, dv$$

et le problème peut maintenant s'écrire :

$$R(u, w) = 0 \quad \forall w(\mathbf{x}) \in H_0^1(\Omega)$$

Pour les systèmes algébriques non linéaires, nous avons utilisé la méthode de Newton qui nécessitait le calcul d'une matrice jacobienne et donc le calcul des dérivées partielles. Revenons maintenant à notre problème non linéaire de départ. Partant d'une première approximation  $u_g^0(\mathbf{x})$  de la solution et vérifiant toutes les conditions aux limites essentielles (homogènes ou non), on cherche une correction  $\delta_u(\mathbf{x})$  telle que :

$$R(u_g^0(\mathbf{x}) + \delta_u(\mathbf{x}), w(\mathbf{x})) = 0 \quad \forall w \in V$$

### ◆ Remarque 8.3

La notation utilisée ici est similaire à celle du cas linéaire. Dans le cas linéaire, on partait d'un relèvement des conditions aux limites essentielles  $u_g$  et on calculait une solution  $\delta_u$  vérifiant les conditions essentielles homogènes de sorte que :

$$u(\mathbf{x}) = u_g(\mathbf{x}) + \delta_u(\mathbf{x})$$

soit la solution du problème. Dans le cas non linéaire, la fonction  $u_g^0$  peut être perçue comme un relèvement des conditions aux limites essentielles et  $\delta_u(\mathbf{x})$  a exactement le même rôle que dans le cas linéaire. ◆

Appliquant le résultat précédent, on trouve :

$$0 = R(u_g^0, w) + D_u(u_g^0, w)[\delta_u] + \|\delta_u\| o(\delta_u, w)$$

Négligeant le terme  $\|\delta_u\| o(\delta_u, w)$ , on doit maintenant résoudre :

$$D_u R(u_g^0, w)[\delta_u] = -R(u_g^0, w)$$

qui est un problème linéaire et qui correspond exactement au système linéaire 8.4 pour les systèmes d'équations algébriques. Si on revient à notre exemple, on a :

$$\begin{aligned}
 D_u R(u_g^0(\mathbf{x}), w(\mathbf{x}))[\delta_u(\mathbf{x})] &= \frac{d}{d\epsilon} [R(u_g^0(\mathbf{x}) + \epsilon\delta_u(\mathbf{x}), w(\mathbf{x}))] \Big|_{\epsilon=0} \\
 &= \frac{d}{d\epsilon} \left( \int_{\Omega} ((u_g^0(\mathbf{x}) + \epsilon\delta_u(\mathbf{x}))^\alpha w(\mathbf{x}) + q(\mathbf{x})\nabla(u_g^0(\mathbf{x}) + \epsilon\delta_u(\mathbf{x})) \cdot \nabla w(\mathbf{x})) dv \right. \\
 &\quad \left. - \int_{\Omega} r(\mathbf{x})w(\mathbf{x}) dv \right) \Big|_{\epsilon=0} \\
 &= \int_{\Omega} (\alpha(u_g^0(\mathbf{x}) + \epsilon\delta_u(\mathbf{x}))^{\alpha-1} \delta_u(\mathbf{x}) w(\mathbf{x}) + q(\mathbf{x})\nabla\delta_u(\mathbf{x}) \cdot \nabla w(\mathbf{x})) dv \Big|_{\epsilon=0} \\
 &= \int_{\Omega} (\alpha(u_g^0(\mathbf{x}))^{\alpha-1} \delta_u(\mathbf{x}) w(\mathbf{x}) + q(\mathbf{x})\nabla\delta_u(\mathbf{x}) \cdot \nabla w(\mathbf{x})) dv
 \end{aligned}$$

Le problème linéarisé consiste donc à :

♠ Pour  $u_g^0(\mathbf{x})$  donné, trouver  $\delta_u(\mathbf{x}) \in H_0^1(\Omega)$  telle que :

$$\int_{\Omega} (\alpha(u_g^0(\mathbf{x}))^{\alpha-1} \delta_u(\mathbf{x}) w(\mathbf{x}) + q(\mathbf{x})\nabla\delta_u(\mathbf{x}) \cdot \nabla w(\mathbf{x})) dv = -R(u_g^0(\mathbf{x}), w(\mathbf{x})) \quad \forall w(\mathbf{x}) \in H_0^1(\Omega)$$

Une fois ce problème linéaire résolu, on obtient une nouvelle approximation de la solution en posant  $u_g^1(\mathbf{x}) = u_g^0(\mathbf{x}) + \delta_u(\mathbf{x})$ . On en arrive à l'algorithme :

- $u_g^0(\mathbf{x})$  étant donné ;
- On résout pour  $k \geq 1$  :

$$\int_{\Omega} (\alpha(u_g^{k-1}(\mathbf{x}))^{\alpha-1} \delta_u(\mathbf{x}) w(\mathbf{x}) + q(\mathbf{x})\nabla\delta_u(\mathbf{x}) \cdot \nabla w(\mathbf{x})) dv = -R((u_g^{k-1}(\mathbf{x})), w(\mathbf{x})) \quad \forall w(\mathbf{x}) \in H_0^1(\Omega) \tag{8.15}$$

- Mise à jour de la solution :  $u_g^k(\mathbf{x}) = u_g^{k-1}(\mathbf{x}) + \delta_u(\mathbf{x})$  ;

**Algorithme 8.2 :** *Méthode de Newton*

1. Étant donné  $\epsilon$ , un critère d'arrêt
2. Étant donné  $N$ , le nombre maximal d'itérations
3. Étant donné  $u_g^0(\mathbf{x})$ , une approximation initiale de la solution du système
4. Résoudre le système linéaire :

$$D_u R(u_g^k, w)[\delta_u] = -R(u_g^k, w) \tag{8.16}$$

5. Mise à jour :

$$u_g^{k+1}(\mathbf{x}) = u_g^k(\mathbf{x}) + \delta_u(\mathbf{x})$$

6. Si  $\frac{\|\delta_u\|_V}{\|u_g^{k+1}\|_V} < \epsilon$ 
  - convergence atteinte
  - écrire la solution  $u_{k+1}(\mathbf{x})$
  - arrêt
7. Si le nombre maximal d'itérations  $N$  est atteint :
  - convergence non atteinte en  $N$  itérations
  - arrêt
8. Retour à l'étape 4

▲

L'algorithme de la méthode de Newton 8.16 converge généralement à l'ordre 2 ce qui signifie que si  $u$  est la solution du problème non linéaire, alors :

$$\|u - u_g^{k+1}\|_V \simeq C \|u - u_g^k\|_V^2$$

et la norme du résidu a exactement le même comportement c.-à-d. :

$$\|R(u_g^{k+1})\|_\infty \simeq C \|R(u_g^k)\|_\infty^2 \quad (8.17)$$

Notre problème non linéaire n'est pas sans rappeler la forme de l'équation 8.7 pour les systèmes algébriques. On peut aussi proposer les algorithmes de points fixes suivant :

- $u_g^0$  étant donné ;
- On résout pour  $k \geq 1$  :

$$\int_{\Omega} q(\mathbf{x}) \nabla u_g^{k+1}(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, dv = - \int_{\Omega} (u_g^k(\mathbf{x}))^\alpha w(\mathbf{x}) \, dv + \int_{\Omega} r(\mathbf{x}) w(\mathbf{x}) \, dv \quad \forall w \in H_0^1(\Omega) \quad (8.18)$$

ou encore :

$$\int_{\Omega} \left( q(\mathbf{x}) \nabla u_g^{k+1} \cdot \nabla w + (u_g^k(\mathbf{x}))^{\alpha-1} u_g^{k+1}(\mathbf{x}) w(\mathbf{x}) \right) \, dv = \int_{\Omega} r(\mathbf{x}) w(\mathbf{x}) \, dv \quad \forall w \in H_0^1(\Omega) \quad (8.19)$$

Les 2 derniers algorithmes sont des variantes de l'algorithme 8.7 pour les systèmes algébriques.

#### ◆ Remarque 8.4

Les algorithmes de points fixes convergent généralement linéairement :

$$\|u - u_g^{k+1}\|_V \simeq C \|u - u_g^k\|_V \quad (C < 1) \quad (8.20)$$

Puisqu'il n'est pas toujours possible de calculer la norme de l'erreur, on note qu'en pratique, la norme du résidu a exactement le même comportement c.-à-d. :

$$\|R(u_g^{k+1})\|_\infty \simeq C \|R(u_g^k)\|_\infty \quad (8.21)$$

Les formes bilinéaires des formulations variationnelles 8.15, 8.18 et 8.19 sont symétriques mais ce n'est pas forcément toujours le cas. La méthode de Newton notamment, mène souvent à des matrices non symétriques. Les matrices symétriques sont plus avantageuses car elles nécessitent moins d'espace mémoire et surtout, si elles sont de plus définies positives, on peut utiliser des méthodes itératives telles le gradient conjugué (préconditionné) pour résoudre les systèmes linéaires correspondants. Dans le cas non symétrique, il existe également des méthodes itératives très efficaces comme la méthode GMRES (« Generalized minimal residual »). ♦

► **Exemple 8.1**

On considère un problème unidimensionnel pour illustrer les notions de convergence des méthodes de Newton et de points fixes. On cherche donc à résoudre dans l'intervalle ]0, 5[ :

$$\begin{cases} -\frac{d^2u}{dx^2}(x) & = \beta e^{u(x)} \\ u(0) = u(5) & = 0 \end{cases}$$

On multiplie par une fonction test  $w$  (en notant dorénavant  $w$  au lieu de  $w(x)$  pour alléger la notation) s'annulant en  $x = 0$  et  $x = 5$  et on intègre sur le domaine. On obtient la formulation variationnelle :

$$\int_0^5 \frac{du}{dx} \frac{dw}{dx} dx = \int_0^5 \beta e^u w dx$$

de sorte que pour une fonction  $u_g^0$  donnée et vérifiant les conditions aux limites essentielles, on définit la fonctionnelle résidu :

$$R(u_g^0, w) = \int_0^5 \left( \frac{du_g^0}{dx} \frac{dw}{dx} - \beta e^{u_g^0} w \right) dx$$

La méthode de Newton appliquée à ce problème nécessite le calcul d'une correction  $\delta_u$  solution de :

$$\int_0^5 \left( \frac{d\delta_u}{dx} \frac{dw}{dx} + \beta e^{u_g^0} \delta_u w \right) dx = - \int_0^5 \left( \frac{du_g^0}{dx} \frac{dw}{dx} - \beta e^{u_g^0} w \right) dx$$

Pour  $\beta = 1/10$  et pour  $u_g^0 = 0$ , les itérations de la méthode de Newton ont donné les résultats suivants :

Méthode de Newton	
Itération	$\ R(u_g^{k+1})\ _\infty$
1	$0,421 \times 10^{+0}$
2	$0,361 \times 10^{-1}$
3	$0,332 \times 10^{-3}$
4	$0,285 \times 10^{-9}$
5	$0,710 \times 10^{-17}$

On constate la convergence quadratique (voir équation 8.17) typique de la méthode. Par contre, une méthode de points fixes possible serait :

$$\int_0^5 \frac{du_g^1}{dx} \frac{dw}{dx} dx = \int_0^5 \beta e^{u_g^0} w dx$$

qui peut aussi s'écrire en posant  $u_g^1 = u_g^0 + \delta_u$  :

$$\int_0^5 \frac{d\delta_u}{dx} \frac{dw}{dx} dx = - \int_0^5 \left( \frac{du_g^0}{dx} \frac{dw}{dx} - \beta e^{u_g^0} w \right) dx$$

On obtient alors les itérés suivants :

Méthode de points fixes			
Itération	$\ R(u_g^{k+1})\ _\infty$	Itération	$\ R(u_g^{k+1})\ _\infty$
1	$0,312 \times 10^{+0}$	7	$0,632 \times 10^{-3}$
2	$0,937 \times 10^{-1}$	8	$0,238 \times 10^{-3}$
3	$0,326 \times 10^{-1}$	$\vdots$	$\vdots$
4	$0,119 \times 10^{-1}$	22	$0,281 \times 10^{-11}$
5	$0,447 \times 10^{-2}$	23	$0,106 \times 10^{-11}$
6	$0,167 \times 10^{-2}$	24	$0,400 \times 10^{-12}$

La convergence est visiblement linéaire (voir l'équation 8.21) avec une constante  $C$  de l'ordre de 0,37. Si on modifie très légèrement en prenant  $\beta$  un peu plus grand ( $\beta = 1/5$ ), la convergence est plus difficile. Partant encore de  $u_g^0 = 0$ , ni la méthode de Newton, ni celle des points fixes n'a convergé. Une stratégie alternative est donc nécessaire. Le problème provient du choix de  $u_g^0 = 0$  qui est trop loin de la solution recherchée. Pour remédier à cela, on peut par exemple résoudre d'abord avec  $\beta = 1/10$  et se servir de la solution obtenue comme point de départ pour la résolution à  $\beta = 1/5$ . On récupère ainsi la convergence quadratique. Une telle stratégie est appelée *méthode de continuation* et il en existe plusieurs variantes plus ou moins sophistiquées. ◀

### 8.4 Exercices

1. On veut résoudre le problème non linéaire suivant par la méthode de Newton :

$$\begin{aligned} -\nabla \cdot [k(1 + \lambda u^2)^m \nabla u(x_1, x_2)] &= f(x_1, x_2) \quad \text{dans } \Omega \\ u(x_1, x_2) &= g(x_1, x_2) \quad \text{sur } \Gamma_0 \\ [k(1 + \lambda u^2)^m] \frac{\partial u}{\partial \mathbf{n}}(x_1, x_2) &= h(x_1, x_2) \quad \text{sur } \Gamma_1 \end{aligned}$$

où  $k$ ,  $\lambda$  et  $m$  sont des constantes positives.

- a) Obtenir la formulation variationnelle, linéariser le problème et donner l'algorithme de la méthode de Newton.
  - b) Pour bien converger, la méthode de Newton requiert une solution de départ vérifiant les conditions essentielles et « pas trop loin » de la solution du problème. Suggérer une façon d'obtenir  $u_0$ .
  - c) Proposer maintenant une méthode de points fixes pour ce problème.
2. On veut résoudre le problème non linéaire suivant par la méthode de Newton :

$$\begin{aligned} -\nabla \cdot (k \nabla u(x_1, x_2)) &= f(x_1, x_2) \quad \text{dans } \Omega \\ u(x_1, x_2) &= g(x_1, x_2) \quad \text{sur } \Gamma_0 \\ k \frac{\partial u}{\partial \mathbf{n}}(x_1, x_2) &= -h(u^4(x_1, x_2) - u_\infty^4) \quad \text{sur } \Gamma_1 \end{aligned}$$

où  $k$ ,  $h$  et  $u_\infty$  sont des constantes positives.

- a) Pour bien converger, la méthode de Newton requiert une solution de départ vérifiant les conditions essentielles et « pas trop loin » de la solution du problème non linéaire. Suggérer une façon d'obtenir  $u_0$ .
- b) Linéariser le problème et donner l'algorithme de la méthode de Newton.



## Chapitre 9

# Problèmes instationnaires

Les problèmes d'évolution (instationnaires) sont d'une très grande importance pratique. La nature même de certains problèmes fait en sorte qu'un état permanent (indépendant du temps) n'est jamais atteint. L'écoulement turbulent d'une rivière en est un exemple. On peut aussi s'intéresser à la transition entre l'état initial d'un système et son état permanent. Dans chaque cas, la solution recherchée varie non seulement en espace mais aussi avec le temps  $t$ .

### 9.1 Rappel sur les équations différentielles ordinaires

Avant de passer au cas des équations aux dérivées partielles, il convient de faire un rappel sur les schémas instationnaires classiques pour les équations différentielles ordinaires. Nous référons le lecteur à Fortin, réf. [17] pour une discussion plus complète. Soit le système d'équations différentielles ordinaires :

$$\begin{cases} \frac{dU}{dt} = f(t, U), & \text{pour } t \in [t_0, t_f] \\ U(t_0) = U_0 \end{cases}$$

On cherchera à obtenir une approximation de la fonction  $u(t)$  pour  $t_0 \leq t \leq t_f$ . Soit  $t_0 < t_1 < \dots < t_N = t_f$  une partition de l'intervalle  $[t_0, t_f]$  et  $\Delta t_{n+1} = t_{n+1} - t_n$ . L'idée la plus simple consiste à remplacer la dérivée temporelle par une formule de différences finies. Par exemple :

$$\begin{cases} \frac{U_{n+1} - U_n}{\Delta t_{n+1}} = f(t_{n+1}, U_{n+1}) \\ U(t_0) = U_0 \end{cases}$$

qui donne le schéma d'Euler implicite d'ordre 1 ou encore :

$$\begin{cases} \frac{3U_{n+1} - 4U_n + U_{n-1}}{2\Delta t} = f(t_{n+1}, U_{n+1}) \\ U(t_0) = U_0 \end{cases}$$

qui est le schéma de différence arrière d'ordre 2 où on a supposé que le pas de temps était constant ( $\Delta t_{n+1} = \Delta t_n = \Delta t$ ). Il s'agit ici d'un schéma à deux pas en ce sens que l'on doit conserver la solution aux deux pas précédents (en  $t = t_n$  et en  $t = t_{n-1}$ ) pour calculer la solution au temps  $t = t_{n+1}$ . On peut aussi introduire un schéma explicite (Euler) :

$$\begin{cases} \frac{U_{n+1} - U_n}{\Delta t_{n+1}} = f(t_n, U_n) \\ U(t_0) = U_0 \end{cases}$$

Plusieurs variantes sont ainsi possibles, chacune ayant ses avantages et inconvénients, la plus importante considération étant la stabilité dont nous discuterons un peu plus loin.

On peut regrouper un certain nombre de ces schémas et introduire une famille de méthodes numériques en considérant la somme pondérée des dérivées de  $U$  de la manière suivante :

$$\theta \frac{dU}{dt}(t_{n+1}) + (1 - \theta) \frac{dU}{dt}(t_n) = \frac{U(t_{n+1}) - U(t_n)}{\Delta t_{n+1}} \quad \text{pour } \theta \in [0, 1] \quad (9.1)$$

On montre facilement à l'aide des développements de Taylor appropriés que :

$$\begin{aligned} & \left( \frac{U(t_{n+1}) - U(t_n)}{\Delta t_{n+1}} \right) - \theta \frac{dU}{dt}(t_{n+1}) - (1 - \theta) \frac{dU}{dt}(t_n) \\ &= \left( \frac{1}{2} - \theta \right) \Delta t_{n+1} \frac{d^2U}{dt^2}(t_n) + \left( \frac{1}{6} - \frac{\theta}{2} \right) (\Delta t_{n+1})^2 \frac{d^3U}{dt^3}(t_n) + \dots \end{aligned}$$

On obtient alors une famille d'approximations paramétrée par  $\theta$  :

- $\theta = 0$  correspond au schéma d'Euler explicite et est d'ordre 1 ;
- $\theta = 1$  correspond au schéma d'Euler implicite et est également d'ordre 1 ;
- $\theta = \frac{1}{2}$  correspond au schéma de Crank-Nicholson d'ordre 2.

Posons  $U_n = U(t_n)$  et  $f_n = f(t_n, U_n)$ , l'équation (9.1) s'écrit :

$$\theta f_{n+1} + (1 - \theta) f_n = \frac{U_{n+1} - U_n}{\Delta t_{n+1}}$$

c.-à-d. :

$$U_{n+1} = U_n + \Delta t_{n+1} [\theta f_{n+1} + (1 - \theta) f_n] \quad (9.2)$$

Dépendant de la valeur de  $\theta$ , on obtient les schémas classiques. Il existe bien entendu d'autres schémas : Runge-Kutta, à pas multiples, etc. mais nous nous limiterons pour le moment au  $\theta$ -schéma.

## 9.2 Formulation quasi-variationnelle

Nous considérerons ici l'équation aux dérivées partielles :

$$\left\{ \begin{array}{l} c(\mathbf{x}, t, u(\mathbf{x}, t)) \frac{\partial u}{\partial t}(\mathbf{x}, t) + p(\mathbf{x}, t, u(\mathbf{x}, t)) u(\mathbf{x}, t) \\ \quad - \nabla \cdot (q(\mathbf{x}, t, u(\mathbf{x}, t)) \nabla u(\mathbf{x}, t)) = r(\mathbf{x}, t, u) \quad \text{dans } \Omega \\ \quad \quad \quad u(\mathbf{x}, t) = g(\mathbf{x}, t) \quad \text{sur } \Gamma_0 \\ \quad \quad \quad q(\mathbf{x}, t, u(\mathbf{x}, t)) \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}, t) = h(\mathbf{x}, t, u(\mathbf{x}, t)) \quad \text{sur } \Gamma_1 \\ \quad \quad \quad u(\mathbf{x}, t_0) = u_0(\mathbf{x}) \end{array} \right. \quad (9.3)$$

que nous avons vue au chapitre 6 mais enrichie d'une dérivée temporelle ainsi que d'une condition initiale. On remarque de plus que les fonctions  $p$ ,  $q$  et  $r$  de même que les conditions aux limites peuvent dépendre du temps (mais pas forcément) et même de la solution  $u$  elle-même dans le cas non linéaire. Il s'agit donc d'un problème d'évolution en temps à partir d'une condition initiale donnée. On cherche ainsi à obtenir une approximation de la fonction  $u(\mathbf{x}, t) = u(x_1, x_2, x_3, t)$  pour  $t_0 \leq t \leq t_f$  et  $\mathbf{x}$  dans un domaine  $\Omega$ . Au temps  $t = t_n$ , multiplions l'équation différentielle par une fonction test  $w$  et intégrons par parties sur l'élément  $K$ , sans prendre de dispositions particulières pour la dérivée temporelle. On obtient :

$$\begin{aligned} \int_K \left( c(\mathbf{x}, t_n, u(\mathbf{x}, t_n)) \frac{\partial u}{\partial t} w(\mathbf{x}) + p(\mathbf{x}, t_n, u(\mathbf{x}, t_n)) u(\mathbf{x}, t) w(\mathbf{x}) + q(\mathbf{x}, t_n, u(\mathbf{x}, t_n)) \nabla u(\mathbf{x}, t_n) \cdot \nabla w \right) dv \\ = \int_K r(\mathbf{x}, t_n, u(\mathbf{x}, t_n)) w(\mathbf{x}) dv + \int_{\partial K \cap \Gamma_1} h(\mathbf{x}, t_n, u(\mathbf{x}, t_n)) w(\mathbf{x}) ds \end{aligned}$$

Cette formulation est dite quasi-variationnelle car la dérivée temporelle n'a pas encore été discrétisée. Nous ne nous en soucierons pas pour l'instant et nous procéderons comme d'habitude. On applique donc la méthode de Ritz sur chaque élément  $K$  mais cette fois en séparant les variables de temps  $t$  et d'espace  $\mathbf{x}$  :

$$u_n = u(\mathbf{x}, t_n)|_K \simeq u^K(\mathbf{x}, t_n) = \sum_{j=1}^{n_d^K} u_j^K(t_n) \psi_j^K(\mathbf{x}) \quad (9.4)$$

Seuls les degrés de liberté dépendent du temps. Les fonctions d'interpolation  $\psi_i^K$  restent les mêmes que dans le cas stationnaire. Mentionnons toutefois l'existence de méthodes d'éléments finis, dites en espace-temps, où les fonctions d'interpolation dépendent explicitement du temps.

Remplaçant dans la formulation quasi-variationnelle, on trouve :

$$\begin{aligned} & \sum_{j=1}^{n_d^K} \dot{u}_j^K(t_n) \int_K c(\mathbf{x}, t_n, u_n) \psi_j^K(\mathbf{x}) w(\mathbf{x}) dv + \\ & \sum_{j=1}^{n_d^K} u_j^K(t_n) \int_K (p(\mathbf{x}, t_n, u_n) \psi_j^K(\mathbf{x}) w(\mathbf{x}) + q(\mathbf{x}, t_n, u_n) \nabla \psi_j^K(\mathbf{x}) \cdot \nabla w) dv \\ & = \int_K r(\mathbf{x}, t_n, u_n) w(\mathbf{x}) dv + \int_{\partial K \cap \Gamma_1} h(\mathbf{x}, t_n, u_n) w(\mathbf{x}) ds \end{aligned}$$

L'expression  $\dot{u}_j^K(t_n)$  désigne la dérivée temporelle du degré de liberté  $u_j^K$  au temps  $t_n$  c.-à-d.  $\frac{du_j^K}{dt}(\mathbf{x}, t_n)$ . Pour obtenir le système élémentaire, on prend successivement  $w(\mathbf{x}) = \psi_i^K(\mathbf{x})$ , pour  $i$  allant de 1 jusqu'à  $n_d^K$ . En notant  $U_n^K$  le vecteur des degrés de liberté de l'élément  $K$  au temps  $t_n$  et  $\dot{U}_n^K$  sa dérivée temporelle, on obtient ainsi le système élémentaire :

$$M^K(U_n^K) \dot{U}_n^K + A^K(U_n^K) U_n^K = F^K(U_n^K) + S^K(U_n^K)$$

où :

$$\begin{aligned} m_{ij,n}^K &= \int_K c(\mathbf{x}, t_n, u_n) \psi_j^K(\mathbf{x}) \psi_i^K(\mathbf{x}) dv \\ a_{ij,n}^K &= \int_K (p(\mathbf{x}, t_n, u_n) \psi_j^K(\mathbf{x}) \psi_i^K(\mathbf{x}) + q(\mathbf{x}, t_n, u_n) \nabla \psi_j^K(\mathbf{x}) \cdot \nabla \psi_i^K(\mathbf{x})) dv \\ f_{i,n}^K &= \int_K r(\mathbf{x}, t_n, u_n) \psi_i^K(\mathbf{x}) dv \\ s_{i,n}^K &= \int_{\partial K \cap \Gamma_1} h(\mathbf{x}, t_n, u_n) \psi_i^K(\mathbf{x}) ds \end{aligned}$$

La notation  $M^K(U_n^K)$  et  $A^K(U_n^K)$  exprime la dépendance de ces matrices sur la solution  $U_n^K$ . Les termes de droite  $F$  et  $S$  prennent en compte respectivement un éventuel terme source et les conditions aux limites naturelles. On pourrait fusionner ces deux termes mais il est mieux avisé de voir explicitement comment sont traités ces conditions aux limites.

Une fois les systèmes élémentaires assemblés, on aura un système global à résoudre à chaque pas de temps de la forme :

$$M(U_n) \dot{U}_n + A(U_n) U_n = F(U_n) + S(U_n) \quad (9.5)$$

#### ◆ Remarque 9.1

Quelques précisions sont nécessaires concernant la notation. Les vecteurs  $F$  et  $S$  ainsi que les matrices  $M$  et  $A$  peuvent dépendre du temps directement ou via la solution  $U_n$ , suivant les fonctions  $r$ ,  $h$ ,  $c$ ,  $p$  et  $q$ . Nous écrirons  $F_n$ ,  $S_n$ ,  $M_n$  et  $A_n$  si ces vecteurs et matrices ne dépendent que du temps (et pas de la solution  $U_n$ ) c.-à-d. si  $r = r(\mathbf{x}, t)$ ,  $h = h(\mathbf{x}, t)$ , etc. Dans le cas non linéaire, nous les noterons  $F(U_n)$ ,  $S(U_n)$ ,  $M(U_n)$  et  $A(U_n)$  (dans le cas où  $r = r(\mathbf{x}, t, u(\mathbf{x}, t))$ ,  $h = h(\mathbf{x}, t, u(\mathbf{x}, t))$ , etc.). Enfin, dans le cas (plus rare) où les vecteurs et/ou les matrices ne dépendent aucunement du temps, les indices seront tout simplement enlevés (correspondant au cas où  $r = r(\mathbf{x})$ ,  $h = h(\mathbf{x})$ , etc.). La matrice  $M(U_n)$  est souvent appelée matrice masse. Si elle dépend de la solution  $u$ , il faut linéariser. ♦

### 9.3 Le theta-schéma

Il existe maintenant plusieurs façons de discrétiser la dérivée temporelle. Supposons maintenant que la matrice  $M(U_n)$  soit inversible. Dans ce cas l'équation 9.5 s'écrit :

$$\begin{cases} \dot{U}_n &= M^{-1}(U_n)[F(U_n) + S(U_n) - A(U_n)U_n] \\ U(0) &= U_0 \end{cases}$$

On peut naturellement appliquer à ce problème le  $\theta$ -schéma 9.2 développé précédemment. Regardons ce qui se passe quand on remplace  $f$  par  $M^{-1}(U)(F(U) + S(U) - A(U)U)$  dans le schéma 9.2 :

$$\begin{aligned} U_{n+1} &= U_n + \Delta t_{n+1} \{ \theta M^{-1}(U_{n+1}) [F(U_{n+1}) + S(U_{n+1}) - A(U_{n+1})U_{n+1}] \\ &\quad + (1 - \theta) M^{-1}(U_n) [F(U_n) + S(U_n) - A(U_n)U_n] \} \end{aligned}$$

Pour simplifier la notation on pose :

$$R_n = M^{-1}(U_n) [F(U_n) + S(U_n) - A(U_n)U_n] \quad \text{ou} \quad M(U_n)R_n = [F(U_n) + S(U_n) - A(U_n)U_n] \quad (9.6)$$

et en multipliant de chaque côté par  $M(U_{n+1})$ , on trouve :

$$\begin{aligned} M(U_{n+1})U_{n+1} &= M(U_{n+1})U_n + \Delta t_{n+1} \{ \theta [F(U_{n+1}) + S(U_{n+1}) - A(U_{n+1})U_{n+1}] \\ &\quad + (1 - \theta) M(U_{n+1})R_n \} \end{aligned}$$

ou encore :

$$\begin{aligned} [M(U_{n+1}) + \theta \Delta t_{n+1} A(U_{n+1})]U_{n+1} &= M(U_{n+1})U_n + \Delta t_{n+1} \{ \theta [F(U_{n+1}) + S(U_{n+1})] \\ &\quad + (1 - \theta) M(U_{n+1})R_n \} \end{aligned} \quad (9.7)$$

Cette expression est *a priori* non linéaire et il est utile de distinguer plusieurs cas.

### 9.3.1 Cas linéaire

Il s'agit du cas le plus simple et le plus classique. Dans le problème type 9.3, les fonctions  $c$ ,  $p$ ,  $q$ ,  $r$  et  $h$  dépendent des variables  $\mathbf{x}$  et  $t$  (mais pas de  $u$ ). On note donc  $M_n$ ,  $A_n$ ,  $F_n$  et  $S_n$ , les vecteurs et matrices du système 9.7. L'équation 9.6 prend la forme :

$$M_n R_n = [F_n + S_n - A_n U_n] \quad (9.8)$$

Notons qu'aucune condition aux limites n'est imposée pour ce problème. On obtient ainsi :

$$[M_{n+1} + \theta \Delta t_{n+1} A_{n+1}] U_{n+1} = M_{n+1} U_n + \Delta t_{n+1} \{ \theta [F_{n+1} + S_{n+1}] + (1 - \theta) M_{n+1} R_n \}$$

On retrouve une expression similaire dans Reddy [31] quoique dans un cas particulier. Comme on a l'habitude de travailler en correction, on pose  $U_{n+1} = U_n + \delta_U$  et on résout à chaque pas de temps :

$$[M_{n+1} + \theta \Delta t_{n+1} A_{n+1}] \delta_U = \Delta t_{n+1} \{ \theta [F_{n+1} + S_{n+1} - A_{n+1} U_n] + (1 - \theta) M_{n+1} R_n \} \quad (9.9)$$

Il en résulte l'algorithme suivant :

1. Pour chaque pas de temps  $n \geq 1$ ,  $U_n$ ,  $F_n$ ,  $S_n$ ,  $M_n$  et  $A_n$  étant connus :
  - a) Résoudre l'équation 9.8 pour  $R_n$  (aucune condition aux limites n'est imposée pour ce problème) ;
  - b) Assembler  $A_{n+1}$ ,  $M_{n+1}$ ,  $F_{n+1}$  et  $S_{n+1}$  ;
  - c) Résoudre l'équation 9.9 pour  $\delta_U$  ;
  - d)  $U_{n+1} = U_n + \delta_U$  ;
2. Passer au pas de temps suivant ;

Les matrices du système doivent être assemblées à tous les pas de temps. Il suffit d'une seule résolution du système 9.9 car il est linéaire.

### 9.3.2 Cas linéaire où la matrice masse est constante

Dans le problème type 9.3, ce cas correspond à la situation où la fonction  $c$  ne dépend que de  $\mathbf{x}$  mais pas de  $t$  ni de  $u$ . Par contre, les autres fonctions  $p$ ,  $q$ , et  $r$  peuvent dépendre de  $\mathbf{x}$  et  $t$ . Les vecteurs et matrices sont notés respectivement  $M$  (constante),  $A_n$ ,  $F_n$  et  $S_n$ . Le dernier terme du système 9.9 se simplifie en utilisant l'équation 9.8. On a ainsi :

$$M_{n+1} R_n = M R_n = M M^{-1} (F_n + S_n - A_n U_n) = F_n + S_n - A_n U_n$$

et par conséquent, l'équation 9.9 devient :

$$[M + \theta \Delta t_{n+1} A_{n+1}] \delta_U = \Delta t_{n+1} \{ \theta [F_{n+1} + S_{n+1} - A_{n+1} U_n] + (1 - \theta) [F_n + S_n - A_n U_n] \} \quad (9.10)$$

Il en résulte l'algorithme simplifié suivant :

1. Pour chaque pas de temps  $n \geq 1$ ,  $U_n, F_n, S_n, M_n$  et  $A_n$  étant connus :
  - a) Assembler  $A_{n+1}, M_{n+1}, F_{n+1}$  et  $S_{n+1}$  ;
  - b) Résoudre l'équation 9.10 pour  $\delta_U$  ;
  - c)  $U_{n+1} = U_n + \delta_U$  ;
2. Passer au pas de temps suivant ;

◆ **Remarque 9.2**

- La matrice  $M$  n'est assemblée qu'une seule fois ;
- Dans le cas où les matrices et vecteurs sont tous constants, de nouvelles simplifications sont possibles et on trouve :

$$[M + \theta \Delta t_{n+1} A] \delta_U = \Delta t_{n+1} [F + S - AU_n] \tag{9.11}$$

- Si le pas de temps est constant, on n'assemble la matrice de gauche qu'une fois pour toutes et elle demeure la même pour tous les pas de temps. Il en résulte une grande économie de temps de calcul ;



### 9.3.3 Cas général non linéaire

Il faut dans ce cas linéariser le système 9.7 et il existe plusieurs possibilités. Notons en premier lieu qu'à un pas de temps donné,  $U_n$  est connu et l'équation 9.6 est toujours linéaire. Une première étape consiste à poser ici encore  $U_{n+1} = U^* + \delta_U$  tout en ne touchant pas aux termes non linéaires. Nous choisirons  $U^*$  comme étant la meilleure approximation en main de  $U_{n+1}$ . Il ne faut toutefois pas oublier de faire évoluer, dans ce vecteur  $U^*$ , les conditions aux limites essentielles pour les imposer au temps  $t = t_{n+1}$ . L'équation 9.7 devient :

$$\begin{aligned}
 [M(U_{n+1}) + \theta \Delta t_{n+1} A(U_{n+1})] \delta_U &= M(U_{n+1}) (U_n - U^*) + \\
 &\Delta t_{n+1} \{ \theta [F(U_{n+1}) + S(U_{n+1}) - A(U_{n+1})U^*] + \\
 &(1 - \theta) M(U_{n+1}) R_n \}
 \end{aligned} \tag{9.12}$$

#### Variante de type points fixes

Une première variante de linéarisation consiste à utiliser un algorithme de points fixes.

1. On se donne un nombre maximum d'itérations  $N$  aussi qu'un critère d'arrêt  $\epsilon_a$  pour les systèmes non linéaires ;
2. Pour chaque pas de temps  $n \geq 1$ ,  $U_n$  étant connu :

- a) Résoudre l'équation 9.6 pour  $R_n$  (aucune condition aux limites n'est imposée pour ce problème) ;
- b) On pose  $U^* = U_n$  et on y impose les conditions aux limites essentielles au temps  $t = t_{n+1}$  ;
- c) Début de la boucle pour la non-linéarité :
  - i. Résoudre :

$$\begin{aligned}
 [M(U^*) + \theta \Delta t_{n+1} A(U^*)] \delta_U &= M(U^*) (U_n - U^*) + \\
 &\Delta t_{n+1} \{ \theta [F(U^*) + S(U^*) - A(U^*)U^*] \\
 &+ (1 - \theta) M(U^*) R_n \} \quad (9.13)
 \end{aligned}$$

ii. Mettre à jour :  $U^* \leftarrow U^* + \delta_U$  ;

iii. Si  $\frac{\|\delta_U\|}{\|U^*\|} \geq \epsilon_a$  et le nombre maximal d'itérations n'est pas atteint, on retourne à 9.13.

d) La solution à ce pas de temps est maintenant connue :  $U_{n+1} = U^*$  ;

3. Passer au pas de temps suivant ;

À chaque pas de temps on doit résoudre le système linéaire pour  $R_n$  sauf pour la méthode d'Euler implicite ( $\theta = 1$ ). On doit également résoudre à chaque pas de temps le système non linéaire pour  $\delta_U$ , ce qui peut devenir coûteux puisqu'à chaque itération non linéaire il faudra assembler les matrices  $A$  et  $M$ . Plusieurs variantes existent et il serait trop long de les énumérer toutes.

### Variante de type Newton

On peut aussi recourir à la méthode de Newton dans le cas de problèmes fortement non linéaires. On revient alors à l'équation 9.12 mais cette fois, on linéarise tous les termes par un développement de Taylor (utilisant la dérivée de Gâteaux). On pose ainsi :

$$\left\{ \begin{array}{l}
 A(U_{n+1}) = A(U^* + \delta_U) \simeq A(U^*) + \frac{\partial A}{\partial U}(U^*) \delta_U \\
 M(U_{n+1}) = M(U^* + \delta_U) \simeq M(U^*) + \frac{\partial M}{\partial U}(U^*) \delta_U \\
 F(U_{n+1}) = F(U^* + \delta_U) \simeq F(U^*) + \frac{\partial F}{\partial U}(U^*) \delta_U \\
 S(U_{n+1}) = S(U^* + \delta_U) \simeq S(U^*) + \frac{\partial S}{\partial U}(U^*) \delta_U
 \end{array} \right.$$

On remplace ensuite ces expressions dans l'équation 9.12. On a ainsi :

$$\begin{aligned} & \left[ M(U^*) + \frac{\partial M}{\partial U}(U^*)\delta_U + \theta\Delta t_{n+1} \left( A(U^*) + \frac{\partial A}{\partial U}(U^*)\delta_U \right) \right] \delta_U \\ &= \left( M(U^*) + \frac{\partial M}{\partial U}(U^*)\delta_U \right) (U_n - U^*) \\ &+ \Delta t_{n+1} \left\{ \theta \left\{ \left( F(U^*) + \frac{\partial F}{\partial U}(U^*)\delta_U \right) + \left( S(U^*) + \frac{\partial S}{\partial U}(U^*)\delta_U \right) - \left( A(U^*) + \frac{\partial A}{\partial U}(U^*)\delta_U \right) U^* \right\} \right. \\ &\left. + (1 - \theta) \left( M(U^*) + \frac{\partial M}{\partial U}(U^*)\delta_U \right) R_n \right\} \end{aligned}$$

Il ne reste plus qu'à négliger les termes d'ordre 2 en  $\delta_U$  et à regrouper à droite tous les termes d'ordre 1. On obtient ainsi :

$$\begin{aligned} & M(U^*)\delta_U - \frac{\partial M}{\partial U}(U^*)\delta_U (U_n - U^*) \\ &+ \theta\Delta t_{n+1} \left\{ A(U^*)\delta_U - \frac{\partial F}{\partial U}(U^*)\delta_U - \frac{\partial S}{\partial U}(U^*)\delta_U + \frac{\partial A}{\partial U}(U^*)\delta_U U^* \right\} - (1 - \theta)\frac{\partial M}{\partial U}(U^*)\delta_U R_n \\ &= M(U^*) (U_n - U^*) + \Delta t_{n+1} \{ \theta \{ F(U^*) + S(U^*) - A(U^*)U^* \} + (1 - \theta)M(U^*)R_n \} \end{aligned} \tag{9.14}$$

L'algorithme résultant est le même que pour la méthode des points fixes mais en remplaçant l'équation 9.13 par 9.14. Si on compare maintenant avec le système 9.13, on remarque que le terme de droite est bien entendu le même mais que la matrice de gauche contient plus de termes. Cette matrice, notée matrice tangente, assure une convergence quadratique. Le prix à payer est cependant assez élevé, à tout le moins dans le cas général. L'expression de la matrice tangente peut paraître très complexe et dans certains cas elle l'est vraiment ! Dans beaucoup de cas simples cependant, la méthode de Newton s'avère tout de même un choix très judicieux.

► **Exemple 9.1**

Considérons l'équation de la chaleur dans le cas où et les propriétés physiques comme la conductivité thermique dépendent de la solution. On a ainsi à résoudre l'équation :

$$\left\{ \begin{array}{ll} \rho c_p(T) \frac{\partial T}{\partial t} - \nabla \cdot (K(T)\nabla u) & = 0 \\ T & = g(\mathbf{x}, t) \text{ sur } \Gamma_0 \\ (K(T)\nabla T) \cdot \mathbf{n} & = h_1(T^4 - T_\infty^4) \text{ sur } \Gamma_1 \\ T(\mathbf{x}, 0) & = T_0(\mathbf{x}) \end{array} \right.$$

La densité est supposée constante ( $\rho = 1000\text{kg/m}^3$ ) et il faut donner un modèle pour tenir compte de la variation des autres coefficients. On prendra par exemple :

$$\begin{aligned} c_p(T) &= 140 + 4,6(T + 273) \text{ J}/(\text{kg } ^\circ\text{C}) \\ K(T) &= 1,46 + 0,0038T^{1,156} \text{ W}/(\text{m}^\circ\text{C}) \end{aligned}$$

La formulation quasi-variationnelle s'écrit :

$$\int_{\Omega} \left( \rho c_p(T) \frac{\partial T}{\partial t} w(\mathbf{x}) + k(T) \nabla T \cdot \nabla w \right) dv = \int_{\Gamma_1} h(T^4 - T_\infty^4) w(\mathbf{x}) ds$$

Dans notre problème modèle 9.3, on a donc :

$$\begin{aligned} c(\mathbf{x}, t, T) &= \rho(140 + 4,6(T + 273)) \text{ de sorte que } \frac{\partial c}{\partial T} = 4,6\rho \\ p(\mathbf{x}, t, T) &= f(\mathbf{x}, t, T) = 0 \\ q(\mathbf{x}, t, T) &= K(T) = 1,46 + 0,0038T^{1,156} \text{ de sorte que } \frac{\partial q}{\partial T} = 0,5928 \times 10^{-2} T^{0,156} \\ h(\mathbf{x}, t, T) &= h_1(T^4 - T_\infty^4) \text{ de sorte que } \frac{\partial h}{\partial T} = 4h_1 T^3 \end{aligned}$$

◀

On remarque que le coefficient de la matrice de masse dépend de la solution  $T$  (et donc du temps). Dans la notation introduite, on a :

$$M(T) = \int_{\Omega} \rho c_p(T) T(\mathbf{x}) w(\mathbf{x}) dv$$

à suivre ....

## 9.4 Un schéma implicite d'ordre 2

Dans cette section, nous introduisons l'un des schémas de résolution de problèmes instationnaires les plus utilisés en pratique, soit le schéma de différences arrière d'ordre 2, aussi connu sous le nom de schéma de Gear.

Au pas de temps  $n + 1$ , l'équation 9.5 s'écrit :

$$\begin{cases} M(U_{n+1}) \dot{U}_{n+1} = [F(U_{n+1}) + S(U_{n+1}) - A(U_{n+1})U_{n+1}] \\ U(0) = U_0 \end{cases}$$

L'idée est alors de remplacer la dérivée temporelle par une différence finie. On montre en effet facilement (voir Fortin [17]) que :

$$\dot{U}_{n+1} = \frac{3U_{n+1} - 4U_n + U_{n-1}}{2\Delta t} + O(\Delta t)^2$$

Le schéma d'ordre 2 est à 2 pas, ce qui pose une petite difficulté au démarrage puisqu'une seule condition initiale n'est connue. Nous verrons plus loin comment contourner cette difficulté. Concentrons nous sur le schéma d'ordre 2 puisque le schéma d'Euler reviendra à la prochaine section. On a maintenant :

$$M(U_{n+1}) \left( \frac{3U_{n+1} - 4U_n + U_{n-1}}{2\Delta t} \right) + A(U_{n+1})U_{n+1} = [F(U_{n+1}) + S(U_{n+1})]$$

Il s'agit, sous sa forme générale d'un problème non linéaire. Bien entendu, si les matrices  $M$  et  $A$  ne dépendent pas de la solution, on retombe sur le cas linéaire. Il faut donc linéariser ce problème. Nous considérons que les solutions aux deux pas de temps précédents  $U_n$  et  $U_{n-1}$  sont connues. On cherche donc à obtenir la solution au temps suivant soit  $U_{n+1}$ . Soit donc  $U^*$  une approximation de  $U_{n+1}$ . Au début d'un nouveau pas de temps, on prend  $U^* = U_n$  par exemple. On pose ensuite  $U_{n+1} = U^* + \delta U$  et on linéarise tous les termes non linéaires. On utilisera bien entendu les développements de Taylor  $M(U^* + \delta U) \simeq M(U^*) + \frac{\partial M}{\partial U} \delta U$  et  $A(U^* + \delta U) \simeq A(U^*) + \frac{\partial A}{\partial U} \delta U$

### 9.4.1 Variante de type points fixes

Il suffit d'évaluer les coefficients de la matrice à l'itération précédente c.-à-d.

$$M(U^*) \left( \frac{3U_{n+1} - 4U_n + U_{n-1}}{2\Delta t} \right) + A(U^*)U_{n+1} = F(U^*) + S(U^*)$$

ou encore :

$$M(U^*) \left( \frac{3U_{n+1}}{2\Delta t} \right) + A(U^*)U_{n+1} = M(U^*) \left( \frac{4U_n - U_{n-1}}{2\Delta t} \right) + F(U^*) + S(U^*)$$

On pose ensuite  $U^* = U_{n+1}$  et on itère jusqu'à convergence c.-à-d. jusqu'à ce que la différence entre deux valeurs successives de  $U_{n+1}$  soit petite.

## 9.5 Résultats numériques

### 9.5.1 Problème thermique

Nous revenons sur un problème déjà rencontré dans le cas stationnaire au chapitre 6 :

$$\rho \frac{\partial c_p T}{\partial t} - \nabla \cdot (\mathbf{K}(\mathbf{x}) \nabla T) = 0$$

où  $\mathbf{K}(\mathbf{x})$  est le tenseur de conductivité thermique. Pour simplifier ce qui suit, on posera simplement que  $\rho = c_p = 1$  et que  $\mathbf{K} = \mathbf{I}$ . Les conditions aux limites sont celles de la figure 6.12 auxquelles on ajoute la condition initiale  $T(\mathbf{x}, 0) = 25^\circ\text{C}$ . On a utilisé un schéma de Gear implicite et un pas de temps adimensionnel  $\Delta t = 1$ .

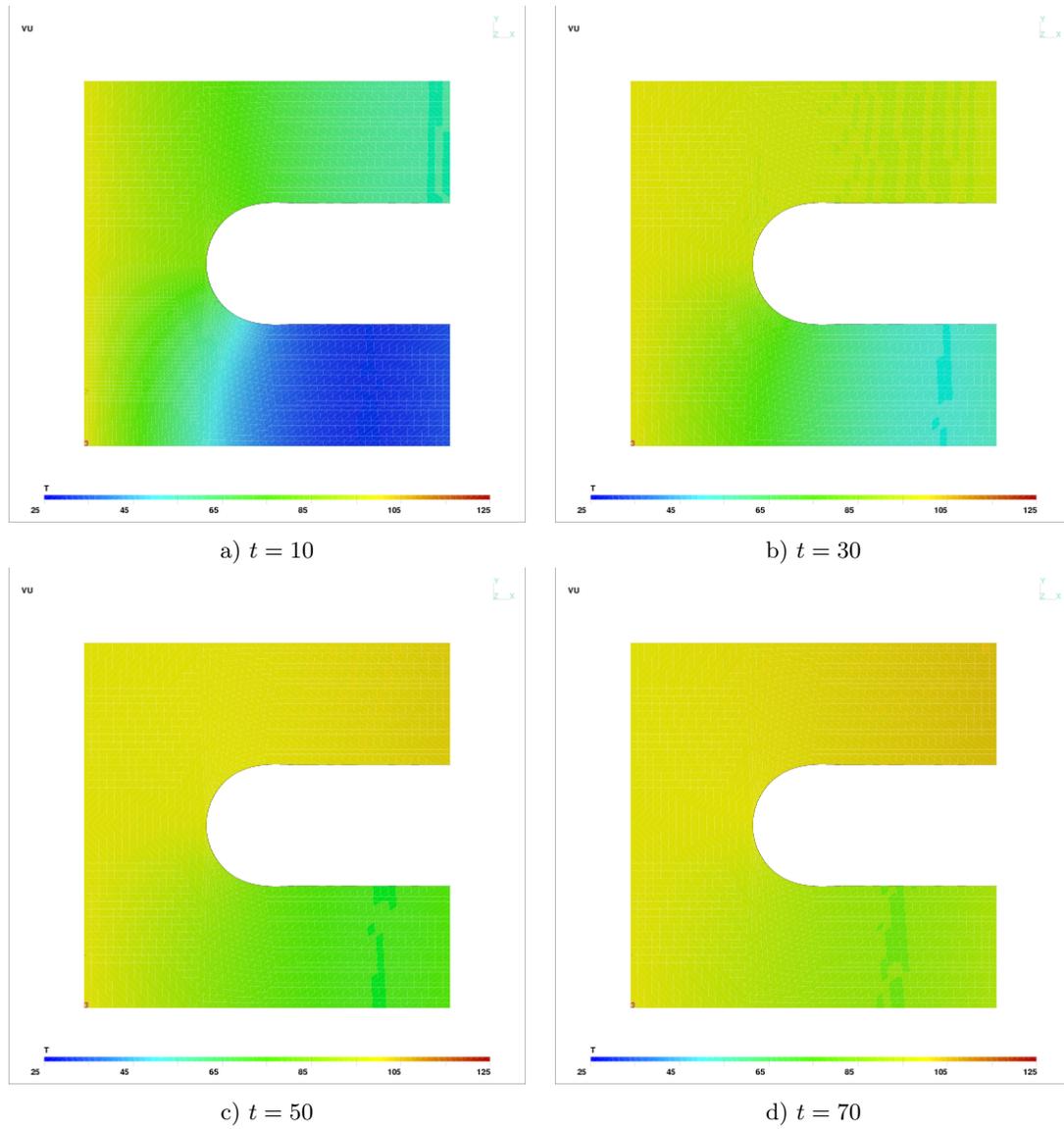


Figure 9.1 – Transfert thermique instationnaire

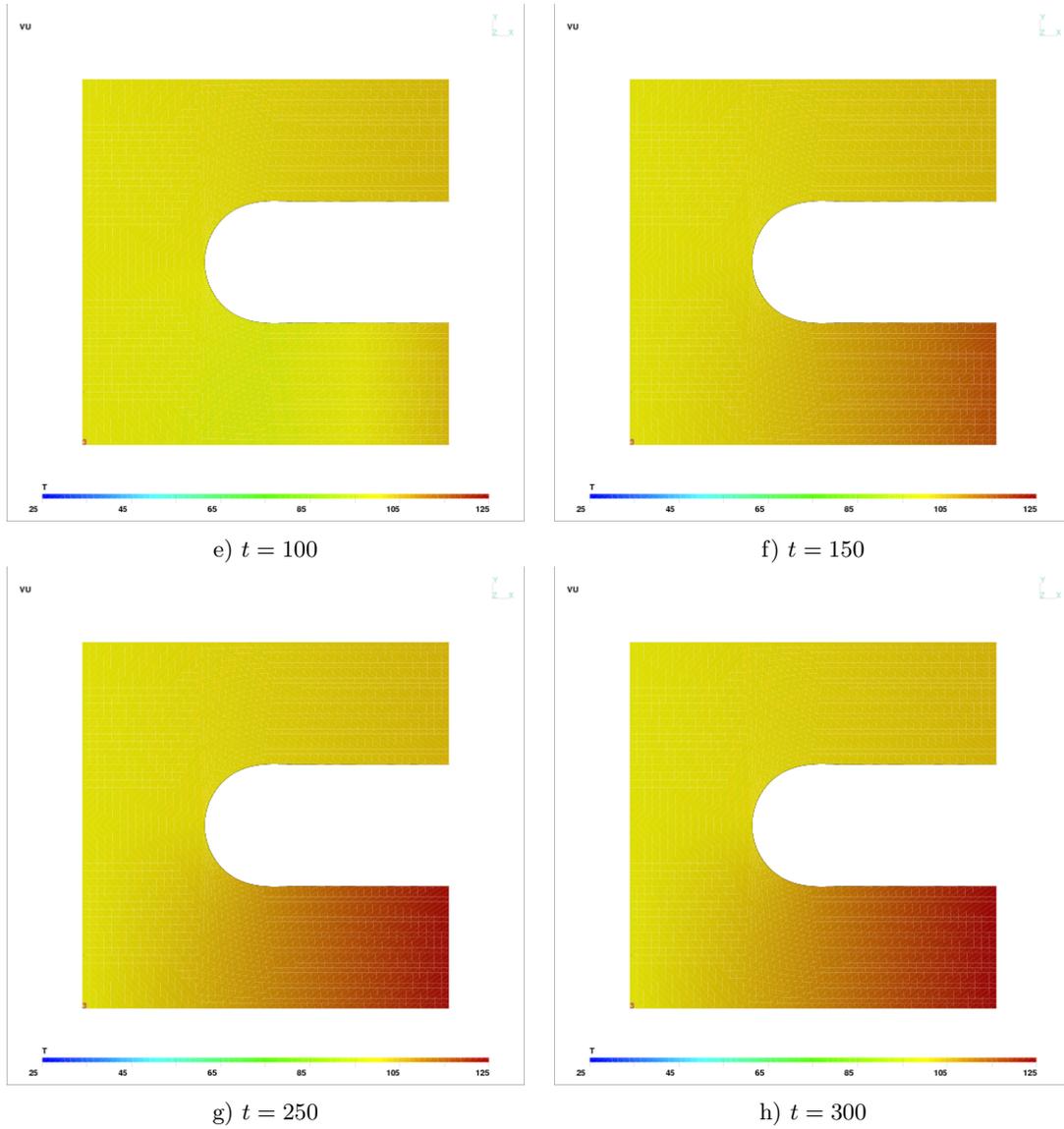


Figure 9.2 – Transfert thermique instationnaire (suite)

On constate l'évolution de la température (Figure 9.1) à partir de la condition initiale à 25°C. La condition aux limites de 100°C fait monter progressivement la température dans tout le domaine, mais nettement plus rapidement dans la moitié supérieure. On constate en effet qu'à  $t = 30$ , la température est d'environ 80°C dans la moitié supérieure. Par la suite ((Figure 9.2), la température se stabilise à un état stationnaire calculé précédemment (voir la figure 6.13).

### 9.5.2 Croissance de populations

On s'intéressera ici à un modèle décrivant l'interaction entre des populations de zooplancton (de nature animale) et de phytoplancton (de nature végétale). Comme nous le verrons, l'interaction est très complexe et donne lieu à des comportements difficiles à prévoir sans une simulation très fine.

Le modèle a été introduit par Medvinsky, Petrovskii, Tikhonova, Malchow et Li en 2002 (voir [26]). C'est un exemple classique d'une équation ou plus précisément d'un système d'équations dites de réaction-diffusion .<sup>1</sup>

$$\begin{cases} \frac{\partial u}{\partial t} - \nabla \cdot (\nabla u) = u(1-u) + \frac{u}{u+h}v \\ \frac{\partial v}{\partial t} - \nabla \cdot (\nabla v) = k\frac{u}{u+h}v - mv \end{cases} \quad (9.15)$$

où  $k = 2$ ,  $m = 0,3$  et  $h = 0,4$  sont des constantes du modèle. Les deux termes de droite (dits de réaction) sont non linéaires et sont d'une importance capitale pour la prédiction des interactions. Des conditions aux limites de type Neumann nulles sont prescrites au bord.

---

1. Nous n'avons pas encore vu comment traiter les systèmes d'équations mais cela n'empêche pas de comprendre la suite.

## Chapitre 10

# Problème de convection-diffusion et stabilisation SUPG

### 10.1 Introduction

La méthode de Galerkin que nous avons largement utilisée jusqu'à maintenant n'est pas sans faille comme nous allons le constater avec l'équation dite de convection diffusion. Pour combler cette lacune, nous allons introduire la méthode de Petrov-Galerkin qui, comme nous l'avons brièvement vu au chapitre 4, consiste à utiliser des fonctions de pondération différentes des fonctions d'interpolation.

L'équation de conservation de l'énergie, en présence de convection, s'écrit (voir Duvaut [14]) :

$$\rho \frac{\partial c_p T}{\partial t} - \nabla \cdot (\mathbf{K}_T(\mathbf{x}) \nabla T) + \mathbf{u} \cdot \nabla T = 0$$

où  $\rho$ ,  $c_p$  et  $\mathbf{K}_T$  sont respectivement la densité, la chaleur spécifique et le tenseur de conductivité thermique du matériau alors que  $\mathbf{u}$  est un vecteur vitesse qui transporte la chaleur. En pratique, le champ  $\mathbf{u}$  est aussi une inconnue du problème et sera, par exemple, obtenu en résolvant les équations de Navier-Stokes (voir le chapitre 12). Pour le moment, nous supposons que  $\mathbf{u}$  est un champ de vecteurs connu. Les applications de cette équation sont nombreuses et variées : problèmes de refroidissement de pièces mécaniques ou électroniques, échauffement du bord d'attaque d'une aile d'avion, etc.

La formulation variationnelle de type Galerkin s'écrit :

$$\int_{\Omega} \rho \frac{\partial c_p T}{\partial t} w + (\mathbf{K}_T(\mathbf{x}) \nabla T) \cdot \nabla w + (\mathbf{u} \cdot \nabla T) w \, dv = 0$$

Cette formulation est bien connue pour pour poser des difficultés numériques lorsque la convection est importante par rapport à la conduction c.-à-d. lorsque  $\|\mathbf{u}\|$  est grand par rapport aux  $K_{T_{ij}}$  du tenseur de conductivité thermique. C'est ce que nous allons maintenant illustrer dans un cas très simple.

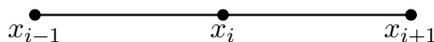


Figure 10.1 – Maillage de deux éléments

On considère l'équation toute simple :

$$-\frac{d}{dx} \left( k \frac{dT}{dx} \right) + u \frac{dT}{dx} = 0 \quad (10.1)$$

où la conductivité thermique est maintenant un scalaire positif et  $u$  une constante que nous supposons positive. On considère le domaine  $[0, 1]$  et des conditions aux limites  $T(0) = 0$  et  $T(1) = 1$  de sorte que l'on vérifie facilement que la solution exacte est :

$$T(x) = \frac{1 - e^{ux/k}}{1 - e^{u/k}}$$

Nous allons faire une brève incursion en différences finies et considérer un maillage uniforme, chaque élément étant de longueur  $h$ . Dans un premier temps, on utilise les différences centrées d'ordre 2 ( $O(h^2)$ ) et un maillage comme celui de la figure 10.1 :

$$\frac{d^2T}{dx^2}(x_i) = \frac{T_{i+1} - 2T_i + T_{i-1}}{h^2} + O(h^2) \text{ ainsi que } \frac{dT}{dx}(x_i) = \frac{T_{i+1} - T_{i-1}}{2h} + O(h^2)$$

de sorte qu'au noeud  $x_i$ , on obtient la récurrence :

$$-\left( \frac{k}{h} + \frac{u}{2} \right) T_{i-1} + \left( \frac{2k}{h} \right) T_i + \left( -\frac{k}{h} + \frac{u}{2} \right) T_{i+1} = 0 \quad (10.2)$$

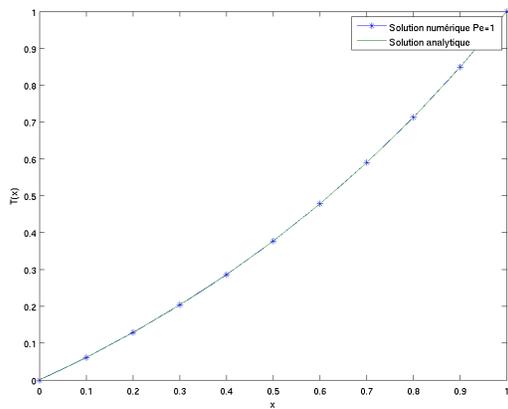
dont les inconnues  $T_i$  sont les valeurs de la température aux noeuds.

Il s'agit d'un système tridiagonal qui peut être résolu très facilement. Les résultats sont illustrés à la figure 10.2. et on remarque rapidement qu'ils sont de plus en plus désastreux au fur et à mesure que  $u$  devient dominant par rapport à  $k$  ou encore que le ratio  $u/k$  augmente. C'est ce que l'on appelle le nombre de Péclet noté  $Pe$ .<sup>1</sup> Il représente le rapport entre le transfert par convection et le transfert par conduction. Il s'agit d'un nombre adimensionnel défini plus précisément par :

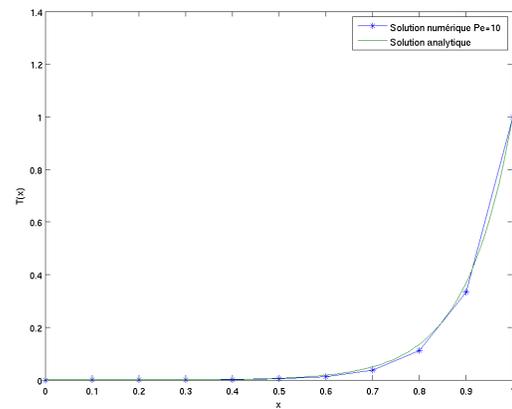
$$Pe = \frac{\rho c_p u L}{k}$$

où  $L$  est une longueur caractéristique du problème, fixée à 1 dans cet exemple tout comme  $c_p$  et  $\rho$ . Plus le nombre de Péclet est grand, plus la convection est dominante et plus les résultats présentent des oscillations. Le cas extrême est celui du transport convectif pur où  $k = 0$ . Physiquement, cela

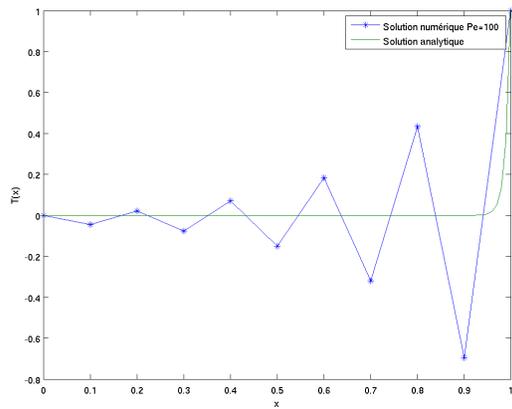
1. Jean Claude Eugène Péclet est un physicien français, né le 10 février 1793 à Besançon, mort à Paris le 6 décembre 1857.



$Pe = 1$



$Pe = 10$



$Pe = 100$

Figure 10.2 – Équation de convection-diffusion : différences centrées et éléments finis linéaires

signifie que l'information vient du sens inverse du vent c.-à-d. en sens inverse du champ de vecteurs  $\mathbf{u}$ . La fumée d'une cheminée suit le vent et il en est de même pour la chaleur qui est convectée par le champ de vitesse c.-à-d. si  $u > 0$ , la chaleur vient de la gauche et inversement si  $u < 0$ .

Devant ces résultats, les spécialistes des différences finies ont eu l'idée d'utiliser une différence arrière pour le terme convectif si  $u > 0$  (avant dans le cas contraire) :

$$\frac{dT}{dx}(x_i) = \frac{T_i - T_{i-1}}{h} + O(h)$$

pour tenir compte du fait que l'information vient principalement de l'amont si  $u$  est grand. On parle alors de différentiation amont («upwinding» en anglais). On obtient la récurrence :

$$-\left(\frac{k}{h} + u\right)T_{i-1} + \left(\frac{2k}{h} + u\right)T_i - \left(\frac{k}{h}\right)T_{i+1} = 0 \quad (10.3)$$

La différence arrière est cependant d'ordre 1 ( $O(h)$ ) seulement et il y donc une perte de précision. On obtient alors les résultats de la figure 10.3. Si les oscillations sont disparues à  $Pe = 100$ , on note tout de même une précision assez faible.

Pour expliquer, au moins partiellement ce qui se passe, on remarque que l'équation 10.3 peut aussi s'écrire, en réarrangeant les termes

$$-\left(k + \frac{uh}{2}\right)\left(\frac{T_{i+1} - 2T_i + T_{i-1}}{h^2}\right) + u\left(\frac{T_{i+1} - T_{i-1}}{2h}\right) = 0 \quad (10.4)$$

et l'on constate que le schéma numérique tend à ajouter un terme de diffusion supplémentaire par rapport à l'équation initiale.

### 10.1.1 Une première approche par éléments finis

Mais comment faire pour obtenir un effet similaire en éléments finis ? Le système élémentaire correspondant à la méthode de Galerkin est :

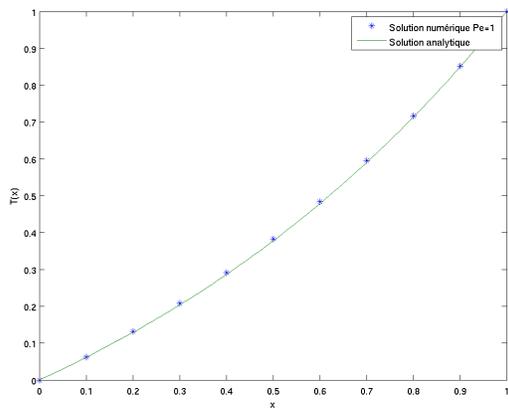
$$\int_{x_1^K}^{x_2^K} k \frac{d\psi_j^K}{dx} \frac{d\psi_i^K}{dx} dx + \int_{x_1^K}^{x_2^K} u \frac{d\psi_j^K}{dx} \psi_i^K dx = 0$$

ou encore sur l'élément de référence :

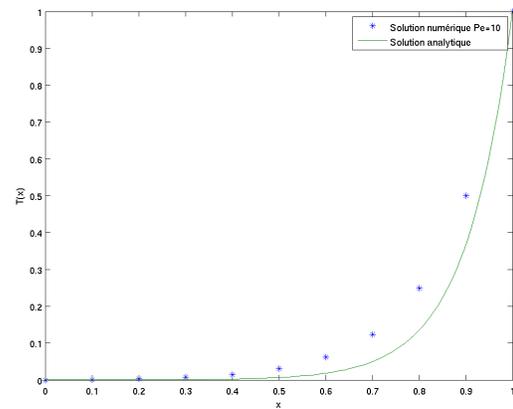
$$\frac{2k}{h} \int_{-1}^1 \frac{d\hat{\psi}_j}{d\xi} \frac{d\hat{\psi}_i}{d\xi} d\xi + u \int_{-1}^1 \frac{d\hat{\psi}_j}{d\xi} \hat{\psi}_i d\xi = 0$$

qui, pour des fonctions d'interpolation linéaires devient :

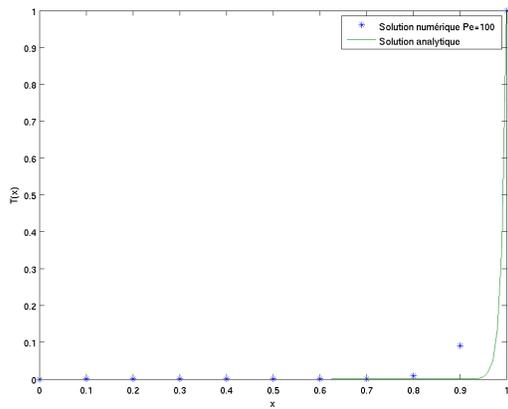
$$\frac{k}{h} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \frac{u}{2} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$



$Pe = 1$



$Pe = 10$



$Pe = 100$

Figure 10.3 – Équation de convection-diffusion : différences arrières

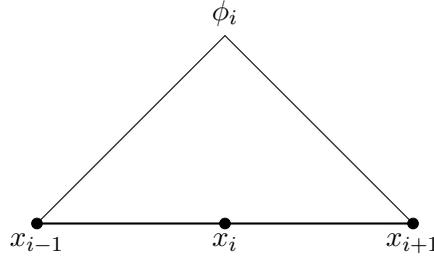


Figure 10.4 – Fonction de Ritz linéaire de type Galerkin

et on considère encore le maillage très simple de la figure 10.1. Une fois la matrice globale assemblée, la deuxième équation, correspondant au noeud  $x_i$ , est exactement l'équation 10.2 et la méthode des éléments finis ne fait pas mieux que celle des différences finies.

Pour obtenir l'effet de la différentiation amont, une solution consiste à modifier les fonctions de pondération dans la formulation éléments finis et de déplacer leur poids vers l'amont. La fonction de Ritz classique linéaire par élément associée au noeud  $x_i$  est illustrée à la figure 10.4. Considérons maintenant la fonction « bulle » sur l'élément de référence :

$$b(\xi) = \frac{3}{4}(1 - \xi^2)$$

et on modifie les fonctions de pondération en les remplaçant par :

$$\tilde{\psi}_1(\xi) = \hat{\psi}_1(\xi) - \alpha b(\xi) = \frac{(1 - \xi)}{2} - \alpha b(\xi)$$

$$\tilde{\psi}_2(\xi) = \hat{\psi}_2(\xi) + \alpha b(\xi) = \frac{(1 + \xi)}{2} + \alpha b(\xi)$$

soit les fonctions habituelles, augmentées de la fonction bulle et d'un paramètre  $\alpha$  à déterminer. La fonction de Ritz du noeud  $x_i$  est illustrée à la figure 10.5 et on constate le biais vers l'amont de la pondération.

Le système élémentaire de type Petrov-Galerkin s'écrit :

$$\frac{2k}{h} \int_{-1}^1 \frac{d\hat{\psi}_j}{d\xi} \frac{d\tilde{\psi}_i}{d\xi} d\xi + u \int_{-1}^1 \frac{d\hat{\psi}_j}{d\xi} \tilde{\psi}_i d\xi = 0$$

et on vérifie facilement que le terme de diffusion reste inchangé. Le système élémentaire de type Petrov-Galerkin devient :

$$\frac{k}{h} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \frac{u}{2} \begin{bmatrix} -(1 - \alpha) & (1 - \alpha) \\ -(1 + \alpha) & (1 + \alpha) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

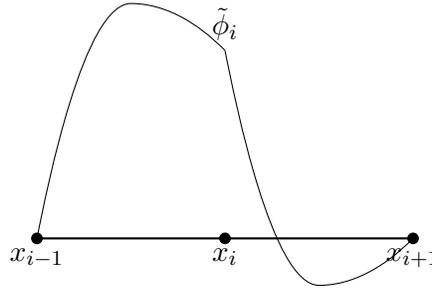


Figure 10.5 – Fonction de Ritz modifiée de type SUPG

et en assemblant sur le maillage de la figure 10.1, on obtient :

$$-\left(\frac{k}{h} + \frac{u}{2}(1 + \alpha)\right) T_{i-1} + \left(\frac{2k}{h} + \alpha u\right) T_i + \left(-\frac{k}{h} + \frac{u}{2}(1 - \alpha)\right) T_{i+1} = 0 \quad (10.5)$$

On remarque immédiatement qu’en posant  $\alpha = 0$ , on retombe sur le schéma de différences centrées. Si  $\alpha = 1$ , on a le schéma de différences arrières et donc nous avons réussi, en modifiant les fonctions de pondération, à obtenir le même effet. Finalement, en posant :

$$\alpha = \coth(\gamma) - \frac{1}{\gamma} \text{ où } \gamma = \frac{uh}{2k} \quad (10.6)$$

on trouve une solution numérique qui est exacte aux noeuds du maillage tel qu’illustré à la figure 10.6. La formule 10.6 donne cette performance que pour le problème simplifié unidimensionnel 10.1. On la retrouve cependant à peu près telle quel dans de nombreux articles en dimension supérieure et pour des équations aux dérivées partielles beaucoup plus générale. Il faut être conscient des limitations d’une telle généralisation...

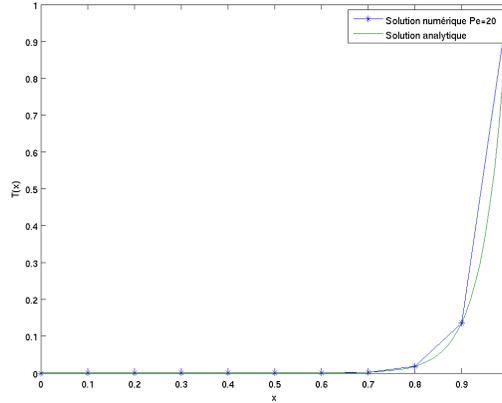
### 10.1.2 Une deuxième approche par éléments finis

La modification des fonctions de pondération semble avoir été la clé du succès de l’approche précédente. On peut arriver à un résultat similaire d’une autre façon. Une approche plus générale consiste à modifier la formulation variationnelle du problème en ajoutant un terme supplémentaire faisant intervenir le résidu (fort) de l’équation à résoudre. Dans le cas de notre problème simplifié, on aurait :

$$\int_0^1 k \frac{dT_h}{dx} \frac{dw}{dx} + u \frac{dT_h}{dx} w \, dx + \sum_K \int_{x_1^K}^{x_2^K} \left( -\frac{d}{dx} \left( k \frac{dT_h}{dx} \right) + u \frac{dT_h}{dx} \right) \tilde{w} \, dx = 0$$

Notons que l’écriture n’est pas arbitraire. Le terme supplémentaire ne peut pas s’écrire :

$$\int_0^1 \left( -\frac{d}{dx} \left( k \frac{dT_h}{dx} \right) + u \frac{dT_h}{dx} \right) \tilde{w} \, dx$$



$$Pe = 20$$

Figure 10.6 – Schéma de type SUPG avec  $\alpha$  défini par 10.6

car ni le terme  $\frac{d}{dx} \left( k \frac{dT_h}{dx} \right)$ , ni les fonctions de pondération  $\tilde{w}$  ne sont continus à la frontière des éléments. La fonction de pondération  $\tilde{w}$  est en effet de la forme :

$$\tilde{w} = \tau_{SUPG} \frac{u}{|u|} \frac{dw}{dx}$$

On peut aussi écrire la formulation variationnelle sous la forme :

$$\sum_K \int_{x_1^K}^{x_2^K} u \frac{dT_h}{dx} (w + \tilde{w})$$

Dans notre problème simplifié,  $\frac{u}{|u|}$  n'est rien d'autre que le signe de  $u$ . Le paramètre  $\tau_{SUPG}$  sera explicité plus tard et dépend de  $k$  et de la taille  $h$  des éléments. De plus, si on prend des fonctions tests  $w$  linéaires,  $\frac{dw}{dx}$  est constant par élément. On a en fait :

$$\frac{d\psi_1^K}{dx} = -\frac{1}{h} \text{ et } \frac{d\psi_2^K}{dx} = \frac{1}{h}$$

de sorte qu'en supposant  $u$  positif et en posant temporairement  $\tau_{SUPG} = :$

$$\psi_1^K + \tau_{SUPG} \frac{u}{|u|} \frac{d\psi_1^K}{dx} = \psi_1^K - \frac{\tau_{SUPG}}{h} \text{ et } \psi_2^K + \tau_{SUPG} \frac{u}{|u|} \frac{d\psi_2^K}{dx} = \psi_2^K + \frac{\tau_{SUPG}}{h}$$

et que l'on retrouve la situation de la figure 10.7 pour la fonction de Ritz attachée au noeud  $x_i$ . Ces fonctions de Ritz sont nettement pondérée vers l'amont mais sont quant à elles discontinues, contrairement à la première approche que nous avons vue.

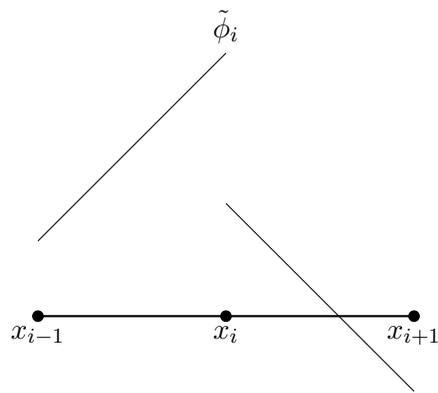


Figure 10.7 – Fonction de Ritz modifiée de type SUPG

## 10.2 Exercices

1. On considère l'équation de convection-diffusion suivante :

$$-\nabla \cdot (\mathbf{K}(\mathbf{x})\nabla T) + \mathbf{u} \cdot \nabla T + c(\mathbf{x})T = f(\mathbf{x}) \text{ dans } \Omega \subset \mathbb{R}^n$$

$$a(\mathbf{x})T + (\mathbf{K}(\mathbf{x})\nabla T) \cdot \mathbf{n} = g(\mathbf{x}) \text{ sur } \Gamma = \partial\Omega$$

où on supposera que :

- $f(\mathbf{x})$  est dans  $L^2(\Omega)$  et  $g(\mathbf{x})$  est dans  $L^2(\Gamma)$  ;
- $c(\mathbf{x})$  est dans  $L^\infty(\Omega)$  et  $a(\mathbf{x})$  est dans  $L^\infty(\Gamma)$  ;
- $\mathbf{u}$  est un champ de vecteurs donné à composantes dans  $W^{1,\infty}(\Omega)$  ;
- $\mathbf{K}$  est un tenseur symétrique dont les composantes sont  $C^1(\bar{\Omega})$  et pour laquelle il existe une constante  $\alpha$  telle que :

$$(\mathbf{K}(\mathbf{x})\boldsymbol{\xi}) \cdot \boldsymbol{\xi} \geq \alpha \|\boldsymbol{\xi}\|_2^2 \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n$$

- Il existe des constantes  $C_1$  et  $C_2$  telles que :

$$c(\mathbf{x}) - \frac{1}{2}\nabla \cdot \mathbf{u} \geq C_1 \geq 0 \text{ presque partout dans } \Omega$$

$$a(\mathbf{x}) + \frac{1}{2}\mathbf{u} \cdot \mathbf{n} \geq C_2 \geq 0 \text{ presque partout sur } \Gamma$$

et l'une au moins des constantes  $C_1$  et  $C_2$  est strictement positive.

Proposer une formulation variationnelle dans un espace fonctionnel approprié et démontrer l'existence et l'unicité de ce problème. Est-ce que l'on peut ramener ce problème à un problème d'optimisation ?

**Suggestion** : Pour la coercivité, on intégrera par parties le terme en  $\mathbf{u} \cdot \nabla T$ . L'identité suivante pourrait aussi être utile :

$$\nabla \cdot (\mathbf{v}w) = (\nabla \cdot \mathbf{v})w + \mathbf{v} \cdot \nabla w$$

2. La discrétisation de problèmes unidimensionnels par éléments finis ou par différences finies mène, dans des cas simples, à des récurrences de la forme :

$$aT_{i+1} + bT_i + cT_{i-1} = 0$$

pour  $i = 1, 2, \dots, n-1$ , auquel on ajoute des conditions aux limites  $T_0 = \alpha_0$  et  $T_n = \alpha_n$ . Si  $b^2 - 4ac > 0$ , montrer que la solution générale de cette récurrence est :

$$T_i = \left( \frac{\alpha_n - \alpha_0 X_2^n}{X_1^n - X_2^n} \right) X_1^i + \left( \frac{\alpha_0 X_1^n - \alpha_n}{X_1^n - X_2^n} \right) X_2^i$$

pour  $X_1$  et  $X_2$  judicieusement choisis.

**Suggestion** : Trouver  $X$  tel que  $T_i = X^i$  soit une solution.

3. En vous servant du résultat du numéro précédent, résoudre les récurrences 10.2, 10.3 et 10.5 et comparer graphiquement avec la solution analytique.

# Chapitre 11

## Application aux problèmes d'élasticité

Tous les problèmes considérés jusqu'à maintenant ne comportaient qu'une seule équation différentielle ou aux dérivées partielles. Il existe cependant plusieurs applications dans lesquelles on doit résoudre un système d'équations. En fait, les systèmes sont de plus en plus la règle plutôt que l'exception.

Nous introduirons ce sujet en discutant des problèmes d'élasticité linéaire, qui ont encore une grande importance pratique. Le problème consiste à calculer les trois composantes du déplacement  $\mathbf{u}(\mathbf{x}) = (u_1(\mathbf{x}), u_2(\mathbf{x}), u_3(\mathbf{x}))$  que subit un corps élastique soumis à diverses sollicitations. On obtient des équations couplées qui ne peuvent être résolues séparément. Il en est de même en mécanique des fluides visqueux incompressibles, où les inconnues sont le champ de vitesse  $(u_1, u_2, u_3)$  et la pression  $p$ . Dans ce dernier cas, on a un système de 4 équations pour ces 4 inconnues que nous étudierons au chapitre suivant.

### 11.1 Problèmes d'élasticité linéaire

On considère un corps fait d'un matériau élastique en dimension 3 (le cas bidimensionnel ne présentant aucune difficulté particulière) et soumis à des forces externes (forces par unité de volume  $\text{N/m}^3$ ) :

$$\mathbf{r}(\mathbf{x}) = \mathbf{r}(x_1, x_2, x_3) = (r_1(x_1, x_2, x_3), r_2(x_1, x_2, x_3), r_3(x_1, x_2, x_3))$$

On souhaite alors déterminer le déplacement  $\mathbf{u}(\mathbf{x})$  d'un point matériel  $\mathbf{x} = (x_1, x_2, x_3)$  occasionné par l'application de ces sollicitations. Pour ce type de problèmes, la notation tensorielle est utile, sinon nécessaire. On retrouvera à l'annexe A quelques rappels sur cette notation et en particulier sur le théorème de la divergence. Introduisant le tenseur (symétrique) des contraintes de Cauchy  $\boldsymbol{\sigma}$ , les équations d'équilibre s'écrivent :

$$-\nabla \cdot \boldsymbol{\sigma} = \mathbf{r}$$

Cette dernière équation est en fait un système de 3 équations :

$$\begin{aligned} - \left( \frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{12}}{\partial x_2} + \frac{\partial \sigma_{13}}{\partial x_3} \right) &= r_1(x_1, x_2, x_3) \\ - \left( \frac{\partial \sigma_{12}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} + \frac{\partial \sigma_{23}}{\partial x_3} \right) &= r_2(x_1, x_2, x_3) \\ - \left( \frac{\partial \sigma_{13}}{\partial x_1} + \frac{\partial \sigma_{23}}{\partial x_2} + \frac{\partial \sigma_{33}}{\partial x_3} \right) &= r_3(x_1, x_2, x_3) \end{aligned}$$

On voit immédiatement l'intérêt d'introduire la notation tensorielle qui est beaucoup plus compacte. Introduisons maintenant deux tenseurs qui seront fort utiles par la suite. Pour un champ de déplacement  $\mathbf{u}$  donné, on définit le *tenseur gradient de déformation* par :

$$\nabla \mathbf{u} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \frac{\partial u_1}{\partial x_3} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \frac{\partial u_2}{\partial x_3} \\ \frac{\partial u_3}{\partial x_1} & \frac{\partial u_3}{\partial x_2} & \frac{\partial u_3}{\partial x_3} \end{bmatrix} \quad (11.1)$$

On définit ensuite le *tenseur de déformation* comme étant la partie symétrique du tenseur gradient de déformation :

$$\gamma(\mathbf{u}) = \frac{\nabla \mathbf{u} + (\nabla \mathbf{u})^T}{2} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \frac{1}{2} \left( \frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} \right) & \frac{1}{2} \left( \frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1} \right) \\ \frac{1}{2} \left( \frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2} \right) & \frac{\partial u_2}{\partial x_2} & \frac{1}{2} \left( \frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2} \right) \\ \frac{1}{2} \left( \frac{\partial u_3}{\partial x_1} + \frac{\partial u_1}{\partial x_3} \right) & \frac{1}{2} \left( \frac{\partial u_3}{\partial x_2} + \frac{\partial u_2}{\partial x_3} \right) & \frac{\partial u_3}{\partial x_3} \end{bmatrix} \quad (11.2)$$

Nous noterons  $\gamma_{ij}$  les différentes composantes du tenseur  $\gamma(\mathbf{u})$ . Notons de plus que le tenseur de déformation n'a pas d'unités.

On multiplie maintenant l'équation d'équilibre par une fonction test vectorielle  $\mathbf{w} = (w_1, w_2, w_3)$  et on intègre par parties sur le domaine  $\Omega$ . On trouve alors, en utilisant le théorème A.7 :

$$\int_{\Omega} \boldsymbol{\sigma} : \nabla \mathbf{w} \, dv - \int_{\Gamma} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \mathbf{w} \, ds = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv$$

Puisque le tenseur  $\boldsymbol{\sigma}$  est symétrique, il est facile de montrer que  $\boldsymbol{\sigma} : \nabla \mathbf{w} = \boldsymbol{\sigma} : \boldsymbol{\gamma}(\mathbf{w})$ , de sorte que la formulation précédente devient :

$$\int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\gamma}(\mathbf{w}) \, dv - \int_{\Gamma} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \mathbf{w} \, ds = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv$$

Il reste à introduire une loi dite *de comportement* reliant le déplacement  $\mathbf{u}$  (ou plus précisément le tenseur  $\boldsymbol{\gamma}(\mathbf{u})$ ) au tenseur des contraintes  $\boldsymbol{\sigma}$ . Pour un matériau dit linéaire élastique, cette relation prend la forme générale :

$$\boldsymbol{\sigma} = \mathcal{C} : \boldsymbol{\gamma}(\mathbf{u}) \quad \text{ou en termes des composantes} \quad \sigma_{ij} = \mathcal{C}_{ijkl} (\boldsymbol{\gamma}(\mathbf{u}))_{kl} = \mathcal{C}_{ijkl} \gamma_{kl} \quad (11.3)$$

où on a utilisé la convention de somme sur les indices répétés. La relation 11.3 n'est qu'une généralisation de la loi de Hooke et  $\mathcal{C}$  est le *tenseur d'élasticité* (du quatrième ordre). On a ainsi une relation linéaire entre les contraintes et les déformations, ce qui justifie l'appellation matériau linéaire élastique. Le tenseur d'élasticité  $\mathcal{C}$  n'est pas tout-à-fait quelconque. On montre en effet qu'il dérive généralement d'un potentiel  $\Phi$  suivant une expression de la forme :

$$\mathcal{C}_{ijkl} = \frac{\partial^2 \Phi}{\partial \gamma_{ij} \partial \gamma_{kl}}$$

Puisque  $\gamma_{ij} = \gamma_{ji}$  et que l'on peut permuter l'ordre de dérivation, on en tire immédiatement les propriétés de symétrie de ce tenseur :

$$\mathcal{C}_{ijkl} = \mathcal{C}_{jikl} = \mathcal{C}_{ijlk} = \mathcal{C}_{klij} \quad (11.4)$$

Le tenseur  $\mathcal{C}$  possède donc au plus 21 composantes différentes et éventuellement non nulles (au lieu des 81 composantes d'un tenseur du quatrième ordre quelconque). Il est courant (et parfois utile) d'exprimer sous forme matricielle l'expression 11.3. Cette relation matricielle est de la forme :

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{bmatrix} = \begin{bmatrix} \tilde{\mathcal{C}}_{11} & \tilde{\mathcal{C}}_{12} & \tilde{\mathcal{C}}_{13} & \tilde{\mathcal{C}}_{14} & \tilde{\mathcal{C}}_{15} & \tilde{\mathcal{C}}_{16} \\ \tilde{\mathcal{C}}_{12} & \tilde{\mathcal{C}}_{22} & \tilde{\mathcal{C}}_{23} & \tilde{\mathcal{C}}_{24} & \tilde{\mathcal{C}}_{25} & \tilde{\mathcal{C}}_{26} \\ \tilde{\mathcal{C}}_{13} & \tilde{\mathcal{C}}_{23} & \tilde{\mathcal{C}}_{33} & \tilde{\mathcal{C}}_{34} & \tilde{\mathcal{C}}_{35} & \tilde{\mathcal{C}}_{36} \\ \tilde{\mathcal{C}}_{14} & \tilde{\mathcal{C}}_{24} & \tilde{\mathcal{C}}_{34} & \tilde{\mathcal{C}}_{44} & \tilde{\mathcal{C}}_{45} & \tilde{\mathcal{C}}_{46} \\ \tilde{\mathcal{C}}_{15} & \tilde{\mathcal{C}}_{25} & \tilde{\mathcal{C}}_{35} & \tilde{\mathcal{C}}_{45} & \tilde{\mathcal{C}}_{55} & \tilde{\mathcal{C}}_{56} \\ \tilde{\mathcal{C}}_{16} & \tilde{\mathcal{C}}_{26} & \tilde{\mathcal{C}}_{36} & \tilde{\mathcal{C}}_{46} & \tilde{\mathcal{C}}_{56} & \tilde{\mathcal{C}}_{66} \end{bmatrix} \begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ \gamma_{33} \\ 2\gamma_{12} \\ 2\gamma_{23} \\ 2\gamma_{13} \end{bmatrix} \quad (11.5)$$

Dans cette relation matricielle, il y a au plus 21 coefficients différents, tout comme pour le tenseur d'élasticité  $\mathcal{C}$  compte tenu de ses symétries. Pour passer de la formulation matricielle 11.5 à la formulation tensorielle 11.3, il suffit de faire les correspondances suivantes. Par exemple, si  $i = j = 1$  dans la relation 11.3, on a :

$$\sigma_{11} = \mathcal{C}_{11kl} \gamma_{kl} = \mathcal{C}_{1111} \gamma_{11} + \mathcal{C}_{1122} \gamma_{22} + \mathcal{C}_{1133} \gamma_{33} + 2\mathcal{C}_{1112} \gamma_{12} + 2\mathcal{C}_{1123} \gamma_{23} + 2\mathcal{C}_{1113} \gamma_{13}$$

tandis que la première équation du système matriciel 11.5 s'écrit :

$$\sigma_{11} = \tilde{C}_{11}\gamma_{11} + \tilde{C}_{12}\gamma_{22} + \tilde{C}_{13}\gamma_{33} + 2\tilde{C}_{14}\gamma_{12} + 2\tilde{C}_{15}\gamma_{23} + 2\tilde{C}_{16}\gamma_{13}$$

de sorte que :

$$C_{1111} = \tilde{C}_{11}, C_{1122} = \tilde{C}_{12}, C_{1133} = \tilde{C}_{13}, C_{1112} = \tilde{C}_{14}, C_{1123} = \tilde{C}_{15}, C_{1113} = \tilde{C}_{16}$$

De même, en prenant  $i = j = 2$  dans 11.3 et la deuxième ligne de 11.5, on conclut que :

$$C_{2211} = C_{1122} = \tilde{C}_{12}, C_{2222} = \tilde{C}_{22}, C_{2233} = \tilde{C}_{23}, C_{2212} = \tilde{C}_{24}, C_{2223} = \tilde{C}_{25}, C_{2213} = \tilde{C}_{26}$$

En parcourant chaque ligne du système matriciel et l'équation correspondante 11.3, on trouve la relation matricielle suivante :

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{bmatrix} = \begin{bmatrix} C_{1111} & C_{1122} & C_{1133} & C_{1112} & C_{1123} & C_{1113} \\ C_{1122} & C_{2222} & C_{2233} & C_{2212} & C_{2223} & C_{2213} \\ C_{1133} & C_{2233} & C_{3333} & C_{3312} & C_{3323} & C_{3313} \\ C_{1112} & C_{2212} & C_{3312} & C_{1212} & C_{1223} & C_{1213} \\ C_{1123} & C_{2223} & C_{3323} & C_{1223} & C_{2323} & C_{2313} \\ C_{1113} & C_{2213} & C_{3313} & C_{1213} & C_{2313} & C_{1313} \end{bmatrix} \begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ \gamma_{33} \\ 2\gamma_{12} \\ 2\gamma_{23} \\ 2\gamma_{13} \end{bmatrix} \quad (11.6)$$

Tous les autres coefficients du tenseur  $\mathbf{C}$  sont obtenus par symétrie.

Que ce soit en utilisant la notation matricielle ou tensorielle, la formulation variationnelle devient alors :

$$\int_{\Omega} (\mathbf{C} : \boldsymbol{\gamma}(\mathbf{u})) : \boldsymbol{\gamma}(\mathbf{w}) \, dv - \int_{\Gamma} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \mathbf{w} \, ds = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv \quad (11.7)$$

Le problème étant d'ordre 2, l'espace de base sera  $H^1(\Omega)$  ou plus précisément  $(H^1(\Omega))^3$  puisque le problème est tridimensionnel. Ce système comporte donc 3 conditions aux limites essentielles à savoir l'imposition des déplacements  $u_i$  sur la frontière ou une partie de celle-ci. Les trois conditions aux limites naturelles correspondantes portent sur le vecteur  $\mathbf{t} = \boldsymbol{\sigma} \cdot \mathbf{n}$  dit de *contraintes normales*. En un point donné de la frontière, on doit imposer 3 conditions aux limites. Dépendant du problème, on impose soit la composante  $u_i$  du déplacement, soit la composante  $t_i$  du vecteur des contraintes normales, mais jamais les deux en même temps.

#### ◆ Remarque 11.1

Le problème variationnel 11.7 est équivalent au problème de minimisation suivant :

$$J(\mathbf{u}) = \inf_{\mathbf{w} \in V} J(\mathbf{w}) = \inf_{\mathbf{w} \in V} \frac{1}{2} \int_{\Omega} (\mathbf{C} : \boldsymbol{\gamma}(\mathbf{w})) : \boldsymbol{\gamma}(\mathbf{w}) \, dv - \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv - \int_{\Gamma} \mathbf{t} \cdot \mathbf{w} \, ds \quad (11.8)$$

où  $V$  est le sous-espace de  $(H^1(\Omega))^3$  des fonctions  $\mathbf{w}$  s'annulant là où on impose des conditions essentielles sur le déplacement  $\mathbf{u}$  ou l'une de ses composantes. Le premier terme correspond à l'énergie élastique de déformation et les deux autres termes correspondent au travail effectué pas la charge imposée. ◆

## 11.2 Matériau linéaire élastique isotrope

Un matériau linéaire est dit isotrope s'il ne possède aucune direction privilégiée ou encore s'il est macroscopiquement homogène. En un point donné, ses propriétés sont les mêmes dans toutes les directions. C'est le cas de la plupart des métaux par exemple.

Dans le cas isotrope, le tenseur d'élasticité  $\mathbf{C}$  est défini par :

$$\mathcal{C}_{ijkl} = \lambda I_{ij} I_{kl} + \mu (I_{ik} I_{jl} + I_{il} I_{jk}) \quad (11.9)$$

où  $\mathbf{I}$  est le tenseur identité ( $I_{ij} = 1$  si  $i = j$ , 0 autrement)<sup>1</sup>. Les paramètres constants  $\lambda$  et  $\mu$  sont les coefficients de Lamé définis par :

$$\mu = \frac{E}{2(1+\nu)} \quad \text{et} \quad \lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad (11.10)$$

où  $E$  est le module d'élasticité (ou module d'Young) et  $\nu$  est le coefficient de Poisson du matériau. Les unités du module de Young sont des pascals (Pa) tandis que le coefficient de Poisson est adimensionnel de sorte que les unités de  $\mu$  et  $\lambda$  sont aussi des pascals. On trouvera dans le tableau suivant quelques valeurs de ces coefficients pour différents matériaux.

Valeurs des coefficients pour des matériaux usuels					
Matériau	$\lambda$ (GPa)	$\mu$ (GPa)	$E$ (GPa)	$\nu$	$K$ (GPa)
Aluminium	60,49	25,93	70	0.35	79
Acier	96,95	76,17	195	0.28	168
Caoutchouc	0,0018			0.5	2.3
Cuivre	130			0.34	139

La relation contraintes-déformations 11.3 prend alors la forme :

$$\sigma_{ij} = (\lambda I_{ij} I_{kl} + \mu (I_{ik} I_{jl} + I_{il} I_{jk})) \gamma_{kl} = \lambda I_{ij} I_{kl} \gamma_{kl} + \mu I_{ik} I_{jl} \gamma_{kl} + \mu I_{il} I_{jk} \gamma_{kl}$$

Puisque  $I_{kl} \gamma_{kl} = \gamma u = \text{tr}(\boldsymbol{\gamma}(\mathbf{u})) = \nabla \cdot \mathbf{u}$ , on a :

$$\sigma_{ij} = \lambda \text{tr}(\boldsymbol{\gamma}(\mathbf{u})) I_{ij} + \mu (I_{ik} \gamma_{kj} + I_{il} \gamma_{jl}) = \lambda \text{tr}(\boldsymbol{\gamma}(\mathbf{u})) I_{ij} + \mu (\gamma_{ij} + \gamma_{ji})$$

de sorte que :

$$\boldsymbol{\sigma} = \lambda (\nabla \cdot \mathbf{u}) \mathbf{I} + 2\mu \boldsymbol{\gamma}(\mathbf{u}) = \left( \frac{E\nu}{(1+\nu)(1-2\nu)} \right) (\nabla \cdot \mathbf{u}) \mathbf{I} + \left( \frac{E}{1+\nu} \right) \boldsymbol{\gamma}(\mathbf{u}) \quad (11.11)$$

1. On retrouve parfois le deuxième terme sous la forme  $2\mu I_{ik} I_{jl}$  qui est incorrecte car alors le tenseur  $\mathbf{C}$  n'a plus les propriétés de symétrie de la relation 11.4.

On peut exprimer la relation 11.11 sous forme matricielle. Il est facile de vérifier que pour un matériau isotrope, l'expression 11.5 devient :

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{bmatrix} = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} \begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ \gamma_{33} \\ 2\gamma_{12} \\ 2\gamma_{23} \\ 2\gamma_{13} \end{bmatrix} \quad (11.12)$$

où on constate qu'il n'y a que 2 coefficients différents ( $E$  et  $\nu$ ). Pour un matériau donné, on peut déterminer expérimentalement ces coefficients. On vérifie alors facilement que les expressions 11.11 et 11.12 sont équivalentes. Nous utiliserons préférablement la notation tensorielle, bien que la formulation matricielle présente aussi certains avantages.

Remplaçant maintenant dans la formulation variationnelle 11.7, on trouve :

$$\int_{\Omega} (\lambda(\nabla \cdot \mathbf{u})\mathbf{I} + 2\mu\boldsymbol{\gamma}(\mathbf{u})) : \boldsymbol{\gamma}(\mathbf{w}) \, dv - \int_{\Gamma} ((\lambda(\nabla \cdot \mathbf{u})\mathbf{I} + 2\mu\boldsymbol{\gamma}(\mathbf{u})) \cdot \mathbf{n}) \cdot \mathbf{w} \, ds = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv$$

ou encore :

$$\int_{\Omega} \lambda(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{w}) + 2\mu\boldsymbol{\gamma}(\mathbf{u}) : \boldsymbol{\gamma}(\mathbf{w}) \, dv = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv + \int_{\Gamma} \mathbf{t} \cdot \mathbf{w} \, ds \quad (11.13)$$

qui est la formulation variationnelle du problème d'élasticité dans le cas d'un matériau isotrope. Dans cette dernière expression, nous avons introduit la condition naturelle sur la contrainte normale :

$$\mathbf{t} = \boldsymbol{\sigma} \cdot \mathbf{n} = \lambda(\nabla \cdot \mathbf{u})\mathbf{n} + 2\mu\boldsymbol{\gamma}(\mathbf{u}) \cdot \mathbf{n} \quad (11.14)$$

Pour démontrer l'existence et l'unicité de la solution du problème 11.13, nous avons besoin d'un résultat préliminaire connu sous le nom d'*inégalité de Korn*.

♣ **Lemme 11.3 (Inégalité de Korn)**

Si  $\mathbf{w} \in (H_{\Gamma_0}^1(\Omega))^3$ , alors il existe une constante  $C(\Omega)$  telle que :

$$\sum_{i,j=1}^3 \|\gamma_{ij}(\mathbf{w})\|_{0,\Omega}^2 \geq C(\Omega)\|\mathbf{w}\|_{1,\Omega}^2 = C(\Omega) (\|w_1\|_{1,\Omega}^2 + \|w_2\|_{1,\Omega}^2 + \|w_3\|_{1,\Omega}^2) \quad (11.15)$$

**Démonstration** : voir Duvaut-Lions [15]. ♣

★ **Théorème 11.1 (Existence et unicité de la solution)**

Le problème variationnel 11.13 possède une solution unique dans  $V = (H^1(\Omega))^3$

**Démonstration :**

Le théorème de lax-Milgram nous assure de l'existence et de l'unicité de la solution si la forme bilinéaire :

$$a(\mathbf{u}, \mathbf{w}) = \int_{\Omega} \lambda(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{w}) + 2\mu\boldsymbol{\gamma}(\mathbf{u}) : \boldsymbol{\gamma}(\mathbf{w}) \, dv$$

est coercive sur  $V$ . Or, en utilisant l'inégalité de Korn, on obtient :

$$a(\mathbf{w}, \mathbf{w}) = \int_{\Omega} \lambda(\nabla \cdot \mathbf{w})^2 + 2\mu\boldsymbol{\gamma}(\mathbf{w}) : \boldsymbol{\gamma}(\mathbf{w}) \, dv \geq \int_{\Omega} 2\mu\boldsymbol{\gamma}(\mathbf{w}) : \boldsymbol{\gamma}(\mathbf{w}) \geq 2\mu C(\Omega) \|\mathbf{w}\|_{1,\Omega}^2$$

ce qui termine la démonstration. ★

### ◆ Remarque 11.2

Le problème variationnel 11.13 est équivalent au problème de minimisation suivant :

$$J(\mathbf{u}) = \inf_{\mathbf{w} \in V} J(\mathbf{w}) = \inf_{\mathbf{w} \in V} \frac{1}{2} \int_{\Omega} \lambda(\nabla \cdot \mathbf{w})^2 + 2\mu\boldsymbol{\gamma}(\mathbf{w}) : \boldsymbol{\gamma}(\mathbf{w}) \, dv - \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv - \int_{\Gamma} \mathbf{t} \cdot \mathbf{w} \, ds \quad (11.16)$$

où  $V$  est le sous-espace de  $(H^1(\Omega))^3$  des fonction  $\mathbf{w}$  s'annulant là où on impose des conditions essentielles sur le déplacement  $\mathbf{u}$  ou l'une de ses composantes. Le premier terme correspond à l'énergie élastique de déformation et les deux autres termes correspondent au travail effectué par la charge imposée. ◆

## 11.3 Matériau linéaire élastique orthotrope

Un matériau est dit orthotrope s'il possède 2 plans de symétrie de comportement. Beaucoup de matériaux composites sont orthotropes. Le bois en est aussi un exemple. En raison de ces plans de symétrie, on peut montrer que la relation contraintes-déformations est de la forme (matricielle) suivante :

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{bmatrix} = \begin{bmatrix} \frac{1-\nu_{23}\nu_{32}}{E_2 E_3 S} & \frac{\nu_{21}+\nu_{31}\nu_{23}}{E_2 E_3 S} & \frac{\nu_{31}+\nu_{21}\nu_{32}}{E_2 E_3 S} & 0 & 0 & 0 \\ \frac{\nu_{12}+\nu_{32}\nu_{13}}{E_1 E_3 S} & \frac{1-\nu_{31}\nu_{13}}{E_1 E_3 S} & \frac{\nu_{32}+\nu_{12}\nu_{31}}{E_1 E_3 S} & 0 & 0 & 0 \\ \frac{\nu_{13}+\nu_{12}\nu_{23}}{E_1 E_2 S} & \frac{\nu_{23}+\nu_{21}\nu_{13}}{E_1 E_2 S} & \frac{1-\nu_{12}\nu_{21}}{E_1 E_2 S} & 0 & 0 & 0 \\ 0 & 0 & 0 & G_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & G_{23} & 0 \\ 0 & 0 & 0 & 0 & 0 & G_{13} \end{bmatrix} \begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ \gamma_{33} \\ 2\gamma_{12} \\ 2\gamma_{23} \\ 2\gamma_{13} \end{bmatrix} \quad (11.17)$$

où :

$$S = \frac{(1 - 2\nu_{21}\nu_{32}\nu_{13} - \nu_{13}\nu_{31} - \nu_{23}\nu_{32} - \nu_{12}\nu_{21})}{E_1 E_2 E_3}$$

À cela s'ajoutent les conditions suivantes, découlant de la symétrie du tenseur :

$$\frac{\nu_{21} + \nu_{31}\nu_{23}}{E_2 E_3 S} = \frac{\nu_{12} + \nu_{32}\nu_{13}}{E_1 E_3 S}, \quad \frac{\nu_{31} + \nu_{21}\nu_{32}}{E_2 E_3 S} = \frac{\nu_{13} + \nu_{12}\nu_{23}}{E_1 E_2 S}, \quad \frac{\nu_{32} + \nu_{12}\nu_{31}}{E_1 E_3 S} = \frac{\nu_{23} + \nu_{21}\nu_{13}}{E_1 E_2 S} \quad (11.18)$$

Il y a donc 9 coefficients indépendants à déterminer (expérimentalement) pour ce type de matériau. Les coefficients  $E_i$  sont les modules d'Young dans chaque direction, les  $\nu_{ij}$  sont les coefficients de Poisson alors que les  $G_{ij}$  sont les modules de cisaillement.

### ◆ Remarque 11.3

Les questions d'existence et d'unicité de la solution sont plus délicates pour un matériau anisotrope. Il faut en effet que la matrice symétrique 11.17 soit de plus définie positive. Il est cependant plus facile de travailler avec la relation inverse de l'équation 11.17 qui s'écrit :

$$\begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ \gamma_{33} \\ 2\gamma_{12} \\ 2\gamma_{23} \\ 2\gamma_{13} \end{bmatrix} = \begin{bmatrix} \frac{1}{E_1} & -\frac{\nu_{21}}{E_2} & -\frac{\nu_{31}}{E_3} & 0 & 0 & 0 \\ -\frac{\nu_{12}}{E_1} & \frac{1}{E_2} & -\frac{\nu_{32}}{E_3} & 0 & 0 & 0 \\ -\frac{\nu_{13}}{E_1} & -\frac{\nu_{23}}{E_2} & \frac{1}{E_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{G_{12}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G_{23}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{13}} \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{bmatrix} \quad (11.19)$$

L'inverse d'une matrice symétrique étant symétrique, on a :

$$\frac{\nu_{21}}{E_2} = \frac{\nu_{12}}{E_1}, \quad \frac{\nu_{31}}{E_3} = \frac{\nu_{13}}{E_1}, \quad \frac{\nu_{32}}{E_3} = \frac{\nu_{23}}{E_2}$$

relations qui sont équivalentes (et beaucoup plus simples) à celles de l'équation 11.18. Dans un premier temps, une matrice définie positive doit avoir des coefficients diagonaux positifs c.à-d.  $E_i > 0$  et  $G_{ij} > 0$ . De plus, l'un des critères pour que cette matrice soit définie positive est que le déterminant de toutes les sous-matrices principales soit positif. On doit alors imposer que :

$$1 - \nu_{12}\nu_{21} > 0, \quad 1 - \nu_{13}\nu_{31} > 0, \quad 1 - \nu_{23}\nu_{32} > 0 \quad (11.20)$$

et enfin que le déterminant de la sous-matrice principale 3 par 3 soit positif c.à-d. que  $S > 0$ . ◆

## 11.4 État plan de contraintes et état plan de déformation

Nous avons présenté le cas général (tridimensionnel) du problème d'élasticité linéaire pour un matériau anisotrope. L'avantage est qu'ainsi, on peut aborder tous les cas. Il existe cependant des situations pour lesquelles, on peut simplifier le problème à deux dimensions. Il faut pour ce faire, déterminer une relation contraintes-déformations qui sera éventuellement différente de l'expression 11.11. Cette nouvelle relation sera ensuite remplacée dans la formulation variationnelle 11.7. Nous nous limiterons de plus au cas orthotrope (comprenant le cas isotrope).

### 11.4.1 État plan de déformation

On considère une section (dans le plan  $x_1x_2$ ) d'un corps élastique (isotrope ou orthotrope) et on suppose que la longueur du corps dans la direction  $x_3$  est grande par rapport aux dimensions de la section. Un exemple typique de cette situation serait un corps cylindrique (de section quelconque) et de grande longueur. On supposera de plus que ce cylindre s'appuie à ses 2 extrémités sur des plans qui préviennent tout déplacement dans la direction  $x_3$ .

Nous pouvons donc supposer que  $u_1 = u_1(x_1, x_2)$  et que  $u_2 = u_2(x_1, x_2)$ . Nous supposons de plus que les déformations ne se produisent que dans le plan  $x_1x_2$  c.-à-d. que  $\gamma_{13} = \gamma_{23} = \gamma_{33} = 0$  ou encore :

$$\frac{\partial u_3}{\partial x_1} = \frac{\partial u_3}{\partial x_2} = \frac{\partial u_3}{\partial x_3} = 0$$

ce qui entraîne que  $u_3 = cte$  et donc que  $u_3 = 0$  puisque le corps est fixé aux deux extrémités.

De l'équation 11.17, on tire que  $\sigma_{13} = \sigma_{23} = 0$  mais, par contre,  $\sigma_{33}$  n'est pas nul :

$$\sigma_{33} = \frac{\nu_{13} + \nu_{12}\nu_{23}}{E_1E_2S} \gamma_{11} + \frac{\nu_{23} + \nu_{21}\nu_{13}}{E_1E_2S} \gamma_{22} \quad (11.21)$$

On se retrouve donc avec un problème purement bidimensionnel à résoudre pour les inconnues en déplacement  $u_1$  et  $u_2$ . La relation contraintes-déformations prend la forme :

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\gamma}(\mathbf{u})$$

ou encore sous forme matricielle :

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} \frac{1-\nu_{23}\nu_{32}}{E_2E_3S} & \frac{\nu_{21}+\nu_{31}\nu_{23}}{E_2E_3S} & 0 \\ \frac{\nu_{12}+\nu_{32}\nu_{13}}{E_1E_3S} & \frac{1-\nu_{31}\nu_{13}}{E_1E_3S} & 0 \\ 0 & 0 & G_{12} \end{bmatrix} \begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ 2\gamma_{12} \end{bmatrix} \quad (11.22)$$

qui, une fois remplacée dans la forme variationnelle 11.7, permet de calculer les déplacements  $u_1$  et  $u_2$ . La contrainte  $\sigma_{33}$  peut ensuite être récupérée par la relation 11.21.

### 11.4.2 État plan de contraintes

Dans ce cas, on considère une plaque mince (d'épaisseur  $h$ ) et on suppose que le plan moyen de cette plaque est dans le plan  $x_1x_2$ . On suppose que les forces volumiques (si elles existent) sont symétriques par rapport au plan moyen et que les éventuelles forces de traction (contrainte normale) au bord de la plaque n'agissent que dans le plan  $x_1x_2$ . Sous ces hypothèses, on peut affirmer que les contraintes agissent uniquement dans le plan  $x_1x_2$  c.-à-d. :

$$\sigma_{13} = \sigma_{23} = \sigma_{33} = 0$$

et que le déplacement  $u_3$  est nul en moyenne sur l'épaisseur de la plaque. On en conclut immédiatement de la relation 11.11 que  $\gamma_{13} = \gamma_{23} = 0$ . De plus, la troisième équation de la relation 11.17 nous donne :

$$0 = \sigma_{33} = \tilde{C}_{31}\gamma_{11} + \tilde{C}_{32}\gamma_{22} + \tilde{C}_{33}\gamma_{33}$$

d'où l'on tire que :

$$\gamma_{33} = \frac{-1}{\tilde{C}_{33}} \left( \tilde{C}_{31}\gamma_{11} + \tilde{C}_{32}\gamma_{22} \right)$$

De la relation 11.17 on tire successivement :

$$\sigma_{11} = \left( \tilde{C}_{11} - \frac{\tilde{C}_{13}\tilde{C}_{31}}{\tilde{C}_{33}} \right) \gamma_{11} + \left( \tilde{C}_{12} - \frac{\tilde{C}_{13}\tilde{C}_{32}}{\tilde{C}_{33}} \right) \gamma_{22}$$

$$\sigma_{22} = \left( \tilde{C}_{21} - \frac{\tilde{C}_{23}\tilde{C}_{31}}{\tilde{C}_{33}} \right) \gamma_{11} + \left( \tilde{C}_{22} - \frac{\tilde{C}_{23}\tilde{C}_{32}}{\tilde{C}_{33}} \right) \gamma_{22}$$

$$\sigma_{12} = \tilde{C}_{44}\gamma_{12}$$

Les relations précédentes peuvent être simplifiées à l'aide des conditions de symétrie 11.20. On obtient ainsi :

$$\sigma_{11} = \left( \frac{E_1}{1-\nu_{12}\nu_{21}} \right) \gamma_{11} + \left( \frac{\nu_{12}E_2}{1-\nu_{12}\nu_{21}} \right) \gamma_{22}$$

$$\sigma_{22} = \left( \frac{\nu_{12}E_2}{1-\nu_{12}\nu_{21}} \right) \gamma_{11} + \left( \frac{E_2}{1-\nu_{12}\nu_{21}} \right) \gamma_{22}$$

$$\sigma_{12} = 2G_{12}\gamma_{12}$$

ou encore, sous forme matricielle :

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} \frac{E_1}{1-\nu_{12}\nu_{21}} & \frac{\nu_{12}E_2}{1-\nu_{12}\nu_{21}} & 0 \\ \frac{\nu_{12}E_2}{1-\nu_{12}\nu_{21}} & \frac{E_2}{1-\nu_{12}\nu_{21}} & 0 \\ 0 & 0 & G_{12} \end{bmatrix} \begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ 2\gamma_{12} \end{bmatrix} \quad (11.23)$$

qui, dans le cas isotrope deviennent :

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} \frac{E}{1-\nu^2} & \frac{\nu E}{1-\nu^2} & 0 \\ \frac{\nu E}{1-\nu^2} & \frac{E}{1-\nu^2} & 0 \\ 0 & 0 & G \end{bmatrix} \begin{bmatrix} \gamma_{11} \\ \gamma_{22} \\ 2\gamma_{12} \end{bmatrix} \quad (11.24)$$

## 11.5 Le maillage

Le maillage pour ce type de problèmes n'a rien de particulier. Chaque élément comprendra des noeuds géométriques et des noeuds de calcul. En chaque noeud de calcul, on associe  $d = 3$  degrés de liberté (2 en dimension 2). Les tables de connectivité des noeuds de calcul *connec*, de numérotation des degrés de liberté *numer* et d'adressage *adres* sont construites de la manière habituelle. Par exemple, pour le problème bidimensionnel ( $d = 2$ ) illustré à la figure 11.1, on a construit un maillage très simple de 8 éléments. Ce maillage n'est présenté que pour illustrer une fois de plus les différents tableaux requis. Le problème sera par la suite résolu sur un maillage beaucoup plus fin et même tridimensionnel.

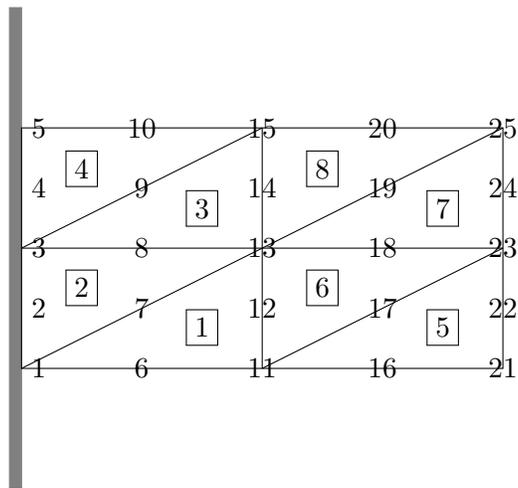


Figure 11.1 – Maillage, numéros d'éléments et numéros noeuds de calcul

On étudie donc une plaque de longueur  $L = 2$  et de hauteur 1 dont l'extrémité gauche est encastree. On choisit une interpolation quadratique et une transformation linéaire vers l'élément de référence telle qu'illustrée à la figure 6.7. On a ainsi  $n_g^K = 3$ ,  $n_c^K = 6$ ,  $n_d^K = 12$  (car nous sommes en dimension 2),  $nel = 8$ ,  $nnoeuds = 25$  et  $nddl = 50$ . Seuls les numéros des éléments et des noeuds sont indiqués sur la figure 11.1. La table de coordonnées des noeuds est de dimension  $nnoeuds \times d$ .

Coordonnées des noeuds					
Table <i>coor</i>					
Noeud	Composante $x_1$	Composante $x_2$	Noeud	Composante $x_1$	Composante $x_2$
1	0,00	0,00	16	1,50	0,00
2	0,00	0,25	17	1,50	0,25
3	0,00	0,50	18	1,50	0,50
4	0,00	0,75	19	1,50	0,75
5	0,00	1,00	20	1,50	1,00
6	0,50	0,00	21	2,00	0,00
7	0,50	0,25	22	2,00	0,25
8	0,50	0,50	23	2,00	0,50
9	0,50	0,75	24	2,00	0,75
10	0,50	1,00	25	2,00	1,00
11	1,00	0,00			
12	1,00	0,25			
13	1,00	0,50			
14	1,00	0,75			
15	1,00	1,00			

La table de connectivité des noeuds est de dimension  $nel \times n_c^K$  et s'écrit :

Noeuds géométriques et de calcul						
Table <i>connec</i>						
Élément	Noeuds géométriques			Autres noeuds		
	Noeud #1	Noeud #2	Noeud #3	Noeud #4	Noeud #5	Noeud #6
1	1	11	13	6	12	7
2	13	3	1	8	2	7
3	3	13	15	8	14	9
4	15	5	3	10	4	9
5	11	21	23	16	22	17
6	23	13	11	18	12	17
7	13	23	25	18	24	19
8	25	15	13	20	14	19

La table de numérotation est de dimension  $nnoeuds \times d$ . Les déplacements des 5 premiers noeuds étant fixés à 0, on omet de les numéroter dans un premier temps, de sorte qu'ils le seront en dernier. Les 2 passages donnent :

Numérotation					
Table <i>numér</i>					
Noeud	$u_1$	$u_2$	Noeud	$u_1$	$u_2$
1	?	?	16	21	22
2	?	?	17	23	24
3	?	?	18	25	26
4	?	?	19	27	28
5	?	?	20	29	30
6	1	2	21	31	32
7	3	4	22	33	34
8	5	6	23	35	36
9	7	8	24	37	38
10	9	10	25	39	40
11	11	12			
12	13	14			
13	15	16			
14	17	18			
15	19	20			

→

Numérotation					
Table <i>numér</i>					
Noeud	$u_1$	$u_2$	Noeud	$u_1$	$u_2$
1	41	42	16	21	22
2	43	44	17	23	24
3	45	46	18	25	26
4	47	48	19	27	28
5	49	50	20	29	30
6	1	2	21	31	32
7	3	4	22	33	34
8	5	6	23	35	36
9	7	8	24	37	38
10	9	10	25	39	40
11	11	12			
12	13	14			
13	15	16			
14	17	18			
15	19	20			

Enfin, la table *adres* est de dimension  $nel \times n_d^K$  et devient :

Adressage Table <i>adres</i>												
Élément	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12
1	41	11	15	1	13	3	42	12	16	2	14	4
2	15	45	41	5	43	3	16	46	42	6	44	4
3	45	15	19	5	17	7	46	16	20	6	18	8
4	19	49	45	9	47	7	20	50	46	10	48	8
5	11	31	35	21	33	23	12	32	36	22	34	24
6	35	15	11	25	13	23	36	16	12	26	14	24
7	15	35	39	25	37	27	16	36	40	26	38	28
8	39	19	15	29	17	27	40	20	16	30	18	28

Remarquons que pour un élément donné, nous avons numéroté d'abord les degrés de liberté associés à la première composante du déplacement et ensuite ceux associés à la deuxième composante. Cette numérotation est parfaitement arbitraire mais une fois ce choix fait, on doit en tenir compte par la suite et particulièrement au moment du calcul du système élémentaire.

## 11.6 Formulation variationnelle élémentaire

Nous nous concentrerons maintenant sur le cas d'un matériau isotrope, plus facile à présenter. Le cas général anisotrope ne pose aucune difficulté conceptuelle supplémentaire.

La formulation variationnelle élémentaire s'obtient en intégrant non plus sur le domaine  $\Omega$  au complet mais sur l'élément  $K$ . On obtient ainsi :

$$\int_K (\lambda(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{w}) + 2\mu\boldsymbol{\gamma}(\mathbf{u}) : \boldsymbol{\gamma}(\mathbf{w})) \, dv = \int_K \mathbf{r}^K \cdot \mathbf{w} \, dv + \int_{\partial K} \mathbf{t}^K \cdot \mathbf{w} \, ds \quad (11.25)$$

où cette fois  $\mathbf{n}^K$  est le vecteur normal à la frontière de l'élément,  $\mathbf{r}^K$  est tout simplement la restriction à l'élément  $K$  de la fonction  $\mathbf{r}(\mathbf{x})$  et  $\mathbf{t}^K$  est la traction définie par l'équation 11.14 exercée à la frontière de l'élément.

### ◆ Remarque 11.4

Dans le cas d'un matériau anisotrope, la formulation variationnelle élémentaire s'écrit tout simplement :

$$\int_K (\mathbf{C} : \boldsymbol{\gamma}(\mathbf{u})) : \boldsymbol{\gamma}(\mathbf{w}) \, dv = \int_K \mathbf{r}^K \cdot \mathbf{w} \, dv + \int_{\partial K} \mathbf{t}^K \cdot \mathbf{w} \, ds$$

où  $\mathbf{C}$  est le tenseur d'élasticité d'ordre 4 associé au matériau considéré ( $\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\gamma}(\mathbf{u})$ ). ◆

## 11.7 Construction des fonctions d'interpolation

La construction des fonctions d'interpolation ne pose aucune difficulté supplémentaire dans le cas des systèmes d'équations. Il existe toutefois la possibilité que les différentes variables du système soient discrétisées par des fonctions d'interpolation de degré différent. Ce n'est pas le cas en élasticité linéaire mais ce le sera au prochain chapitre pour le problème de Stokes.

Pour obtenir le système élémentaire, on pose :

$$\mathbf{u}(\mathbf{x})|_K \simeq \mathbf{u}^K(\mathbf{x}) = \sum_{j=1}^{n_d^K} \alpha_j^K \Psi_j^K(\mathbf{x})$$

On remarque la notation quelque peu particulière. En premier lieu, les fonctions tests  $\Psi_j^K(\mathbf{x})$  sont vectorielles. Si on suppose que l'on a  $n_c^K$  noeuds de calcul et donc  $n_d^K = d n_c^K$  degrés de liberté, le vecteur des inconnues élémentaires devient :

$$\alpha^K = \begin{bmatrix} \alpha_1^K \\ \alpha_2^K \\ \vdots \\ \alpha_{n_c^K}^K \\ \alpha_{n_c^K+1}^K \\ \alpha_{n_c^K+2}^K \\ \vdots \\ \alpha_{2n_c^K}^K \\ \alpha_{2n_c^K+1}^K \\ \alpha_{2n_c^K+2}^K \\ \vdots \\ \alpha_{3n_c^K}^K \end{bmatrix} = \begin{bmatrix} u_{1,1}^K \\ u_{1,2}^K \\ \vdots \\ u_{1,n_c^K}^K \\ u_{2,1}^K \\ u_{2,2}^K \\ \vdots \\ u_{2,n_c^K}^K \\ u_{3,1}^K \\ u_{3,2}^K \\ \vdots \\ u_{3,n_c^K}^K \end{bmatrix} = \begin{bmatrix} U_1^K \\ U_2^K \\ U_3^K \end{bmatrix} \quad (11.26)$$

et contient alors tous les degrés de liberté de l'élément. Tout comme dans la table d'adressage, nous avons numéroté en premier tous les degrés de liberté associés à la première composante du déplacement, ensuite ceux associés à la deuxième composante et ainsi de suite. D'autres numérotations sont bien entendu envisageables.

Si on note  $\psi_j^K$  les  $n_c^K$  fonctions d'interpolation attachées aux  $n_c^K$  noeuds de calcul, on exprime

alors les  $3n_c^K$  fonctions tests  $\Psi_j^K$  sous la forme :

$$\Psi^K = \begin{bmatrix} \Psi_1^K \\ \Psi_2^K \\ \vdots \\ \Psi_{n_c^K}^K \\ \hline \Psi_{n_c^K+1}^K \\ \Psi_{n_c^K+2}^K \\ \vdots \\ \Psi_{2n_c^K}^K \\ \hline \Psi_{2n_c^K+1}^K \\ \Psi_{2n_c^K+2}^K \\ \vdots \\ \Psi_{3n_c^K}^K \end{bmatrix} = \begin{bmatrix} (\psi_1^K, 0, 0) \\ (\psi_2^K, 0, 0) \\ \vdots \\ (\psi_{n_c^K}^K, 0, 0) \\ \hline (0, \psi_1^K, 0) \\ (0, \psi_2^K, 0) \\ \vdots \\ (0, \psi_{n_c^K}^K, 0) \\ \hline (0, 0, \psi_1^K) \\ (0, 0, \psi_2^K) \\ \vdots \\ (0, 0, \psi_{n_c^K}^K) \end{bmatrix} \quad (11.27)$$

de sorte que :

$$\sum_{j=1}^{n_d^K} \alpha_j^K \Psi_j^K(\mathbf{x}) = \left( \sum_{j=1}^{n_c^K} u_{1,j}^K \psi_j^K(\mathbf{x}), \sum_{j=1}^{n_c^K} u_{2,j}^K \psi_j^K(\mathbf{x}), \sum_{j=1}^{n_c^K} u_{3,j}^K \psi_j^K(\mathbf{x}) \right) = (u_1^K(\mathbf{x}), u_2^K(\mathbf{x}), u_3^K(\mathbf{x}))$$

On obtient alors le *système élémentaire* de dimension  $n_d^K$  suivant :

$$A^K \alpha^K = F^K + S^K \quad (11.28)$$

où :

$$a_{ij}^K = \int_K (\lambda(\nabla \cdot \Psi_j^K)(\nabla \cdot \Psi_i^K) + 2\mu \gamma(\Psi_j^K) : \gamma(\Psi_i^K)) dv$$

$$f_i^K = \int_K \mathbf{r}^K \cdot \Psi_i^K dv$$

$$s_i^K = \int_{\partial K} \mathbf{t}^K \cdot \Psi_i^K ds$$

Le calcul du système élémentaire se divise naturellement en plusieurs blocs (de dimension  $n_c$ ), selon la forme particulière des fonctions tests  $\Psi_i^K$  et  $\Psi_j^K$ . On peut en effet décomposer le système sous la forme :

$$\begin{bmatrix} A_{11}^K & A_{12}^K & A_{13}^K \\ A_{21}^K & A_{22}^K & A_{23}^K \\ A_{31}^K & A_{32}^K & A_{33}^K \end{bmatrix} \begin{bmatrix} U_1^K \\ U_2^K \\ U_3^K \end{bmatrix} = \begin{bmatrix} F_1^K \\ F_2^K \\ F_3^K \end{bmatrix} + \begin{bmatrix} S_1^K \\ S_2^K \\ S_3^K \end{bmatrix}$$

Les matrices  $A_{ij}^K$  sont de dimension  $n_c^K$  par  $n_c^K$ . Par exemple, la matrice  $A_{11}^K$  fait intervenir les fonctions  $(\psi_j^K, 0, 0)$  et  $(\psi_i^K, 0, 0)$  de sorte que :

$$(A_{11}^K)_{ij} = \int_K \left( \lambda \frac{\partial \psi_j^K}{\partial x_1} \frac{\partial \psi_i^K}{\partial x_1} + 2\mu \left( \frac{\partial \psi_j^K}{\partial x_1} \frac{\partial \psi_i^K}{\partial x_1} + \frac{1}{2} \frac{\partial \psi_j^K}{\partial x_2} \frac{\partial \psi_i^K}{\partial x_2} + \frac{1}{2} \frac{\partial \psi_j^K}{\partial x_3} \frac{\partial \psi_i^K}{\partial x_3} \right) \right) dv$$

De même :

$$(A_{12}^K)_{ij} = \int_K \left( \lambda \frac{\partial \psi_j^K}{\partial x_2} \frac{\partial \psi_i^K}{\partial x_1} + \mu \left( \frac{\partial \psi_j^K}{\partial x_1} \frac{\partial \psi_i^K}{\partial x_2} \right) \right) dv = (A_{21}^K)_{ji}$$

$$(A_{13}^K)_{ij} = \int_K \left( \lambda \frac{\partial \psi_j^K}{\partial x_3} \frac{\partial \psi_i^K}{\partial x_1} + \mu \left( \frac{\partial \psi_j^K}{\partial x_1} \frac{\partial \psi_i^K}{\partial x_3} \right) \right) dv = (A_{31}^K)_{ji}$$

$$(A_{22}^K)_{ij} = \int_K \left( \lambda \frac{\partial \psi_j^K}{\partial x_2} \frac{\partial \psi_i^K}{\partial x_2} + 2\mu \left( \frac{1}{2} \frac{\partial \psi_j^K}{\partial x_1} \frac{\partial \psi_i^K}{\partial x_1} + \frac{\partial \psi_j^K}{\partial x_2} \frac{\partial \psi_i^K}{\partial x_2} + \frac{1}{2} \frac{\partial \psi_j^K}{\partial x_3} \frac{\partial \psi_i^K}{\partial x_3} \right) \right) dv$$

$$(A_{23}^K)_{ij} = \int_K \left( \lambda \frac{\partial \psi_j^K}{\partial x_3} \frac{\partial \psi_i^K}{\partial x_2} + \mu \left( \frac{\partial \psi_j^K}{\partial x_2} \frac{\partial \psi_i^K}{\partial x_3} \right) \right) dv = (A_{32}^K)_{ji}$$

$$(A_{33}^K)_{ij} = \int_K \left( \lambda \frac{\partial \psi_j^K}{\partial x_3} \frac{\partial \psi_i^K}{\partial x_3} + 2\mu \left( \frac{1}{2} \frac{\partial \psi_j^K}{\partial x_1} \frac{\partial \psi_i^K}{\partial x_1} + \frac{1}{2} \frac{\partial \psi_j^K}{\partial x_2} \frac{\partial \psi_i^K}{\partial x_2} + \frac{\partial \psi_j^K}{\partial x_3} \frac{\partial \psi_i^K}{\partial x_3} \right) \right) dv$$

et enfin :

$$(F_1^K)_i = \int_K r_1^K \psi_i^K dv, \quad (S_1^K)_i = \int_{\partial K} t_1^K \psi_i^K ds$$

$$(F_2^K)_i = \int_K r_2^K \psi_i^K dv, \quad (S_2^K)_i = \int_{\partial K} t_2^K \psi_i^K ds$$

$$(F_3^K)_i = \int_K r_3^K \psi_i^K dv, \quad (S_3^K)_i = \int_{\partial K} t_3^K \psi_i^K ds$$

## 11.8 Passage à l'élément de référence

Le calcul de tous ces termes est une fois de plus effectué après le passage à l'élément de référence. Cela suppose *a priori* le choix d'un type d'éléments : triangles, quadrilatères, tétraèdres, etc. Le choix du type d'éléments détermine, comme au chapitre 6 une transformation  $T^K$  vérifiant :

$$T^K(\boldsymbol{\xi}_i) = \mathbf{x}_i^K \quad \text{ou inversement} \quad (T^K)^{-1}(\mathbf{x}_i^K) = \boldsymbol{\xi}_i$$

pour chacun des  $n_g^K$  noeuds géométriques de l'élément. On transforme les dérivées partielles de la manière habituelle à l'aide de la relation 6.6. On obtient ainsi par exemple :

$$\frac{\partial \psi_j^K}{\partial x_1} = \frac{\partial \hat{\psi}_j}{\partial \xi} B_{11}^K + \frac{\partial \hat{\psi}_j}{\partial \eta} B_{12}^K + \frac{\partial \hat{\psi}_j}{\partial \zeta} B_{13}^K$$

Rappelons que la matrice  $B^K$  et la matrice jacobienne  $DT^K$  sont reliées par l'équation  $B^K = (DT^K)^{-t}$ . Le système élémentaire devient alors :

$$\begin{aligned} (A_{11}^K)_{ij} &= \int_{\hat{K}} \left[ (\lambda + 2\mu) \left( \frac{\partial \hat{\psi}_j}{\partial \xi} B_{11}^K + \frac{\partial \hat{\psi}_j}{\partial \eta} B_{12}^K + \frac{\partial \hat{\psi}_j}{\partial \zeta} B_{13}^K \right) \left( \frac{\partial \hat{\psi}_i}{\partial \xi} B_{11}^K + \frac{\partial \hat{\psi}_i}{\partial \eta} B_{12}^K + \frac{\partial \hat{\psi}_i}{\partial \zeta} B_{13}^K \right) \right. \\ &+ \mu \left( \frac{\partial \hat{\psi}_j}{\partial \xi} B_{21}^K + \frac{\partial \hat{\psi}_j}{\partial \eta} B_{22}^K + \frac{\partial \hat{\psi}_j}{\partial \zeta} B_{23}^K \right) \left( \frac{\partial \hat{\psi}_i}{\partial \xi} B_{21}^K + \frac{\partial \hat{\psi}_i}{\partial \eta} B_{22}^K + \frac{\partial \hat{\psi}_i}{\partial \zeta} B_{23}^K \right) \\ &+ \left. \mu \left( \frac{\partial \hat{\psi}_j}{\partial \xi} B_{31}^K + \frac{\partial \hat{\psi}_j}{\partial \eta} B_{32}^K + \frac{\partial \hat{\psi}_j}{\partial \zeta} B_{33}^K \right) \left( \frac{\partial \hat{\psi}_i}{\partial \xi} B_{31}^K + \frac{\partial \hat{\psi}_i}{\partial \eta} B_{32}^K + \frac{\partial \hat{\psi}_i}{\partial \zeta} B_{33}^K \right) \right] J^K d\hat{v} \end{aligned} \quad (11.29)$$

et des expressions similaires pour les autres coefficients. Malgré l'aspect rébarbatif de ces expressions, il faut se rappeler que toutes les dérivées partielles ne sont évaluées qu'une seule fois (aux points d'intégration). C'est l'avantage de la technique du passage à l'élément de référence. Les coefficients  $B_{ij}^K$  dépendent de l'élément mais ne représentent pas un effort de calcul considérable.

## 11.9 Évaluation du système élémentaire

Ici encore, l'intégration numérique sera le plus souvent utilisée. La diversité des expressions à intégrer de même que leur complexité rendent très important le choix de la formule de quadrature. Dans la mesure du possible on choisira le nombre de points d'intégrations de sorte que le système élémentaire soit intégré exactement. Malheureusement ce n'est pas toujours possible. Il faut aussi s'attendre à ce que cette étape soit plus coûteuse qu'auparavant en raison du nombre plus élevé d'intégrales à évaluer et du plus grand nombre de points d'intégration nécessaires pour les évaluer convenablement.

Supposons par exemple que l'on choisisse une interpolation quadratique sur des tétraèdres et une transformation géométrique linéaire (illustrée à la figure 6.6) vers l'élément de référence. Les fonctions d'interpolation  $\hat{\psi}_i(\xi)$  sont quadratiques et le jacobien  $J^K$  est constant par élément. Si on regarde l'expression 11.29, on constate que puisque les coefficients  $B_{ij}^K$  sont également constants par élément et que les dérivées partielles des fonctions d'interpolation sont linéaires, l'intégrand est de degré maximum 2. La table D.3 nous indique alors qu'une formule de Hammer à 4 points serait suffisante pour évaluer la matrice élémentaire.

Une nouvelle analyse est nécessaire pour choisir la formule de quadrature pour le terme de droite qui dépend d'une fonction  $\mathbf{r}$ .

## 11.10 Assemblage et imposition des conditions aux limites

On assemble le système global exactement comme dans les chapitres précédents. Les systèmes élémentaires sont de taille plus imposantes mais la technique reste la même. Le système linéaire global résultant pourra, surtout dans le cas tridimensionnel, devenir impressionnant et même dans certains cas, prohibitif en raison de l'espace-mémoire nécessaire.

Pour imposer les conditions aux limites, on applique la technique déjà connue en se rapportant aux équations 5.20 et 5.21.

## 11.11 Résolution du système global

Ici encore, la seule différence réside dans la taille des systèmes linéaires obtenus. Si pour la majorité des problèmes bidimensionnels, on peut utiliser une décomposition  $LU$  avec un stockage bien avisé de la matrice, ce n'est plus le cas en dimension 3. On doit alors avoir recours aux méthodes itératives de résolution qui favorisent un stockage plus efficace de la matrice. Les méthodes itératives sont cependant encore peu utilisées en pratique.

## 11.12 Visualisation des résultats

La visualisation d'une solution numérique n'est pas du tout une mince tâche, particulièrement dans le cas tridimensionnel. On a recours à des logiciels spécialisés comme  $VU$  [27].

## 11.13 Applications

### 11.13.1 Essai en cisaillement

Nous considérerons le problème d'une poutre fixée à son extrémité gauche. On impose un cisaillement sur la face droite de la tige c'est-à-dire  $\boldsymbol{\sigma} \cdot \mathbf{n} = (0, -\sigma_0, 0)$ . On a construit le maillage tridimensionnel de la figure 11.2 consistant en 750 éléments tétraédriques quadratiques pour un total de 3663 degrés de liberté. On remarque immédiatement que pour un problème tridimensionnel, le nombre de degrés de liberté augmente très rapidement. Il ne faut pas perdre de vue qu'un maillage de 750 éléments est nettement insuffisant pour la plupart des applications.

La figure 11.3 montre clairement la déformation due au cisaillement imposé. On y a déformé le maillage d'une quantité proportionnelle au déplacement calculé  $(u_1, u_2, u_3)$ .

### 11.13.2 Plaque trouée en élongation

On considère la plaque trouée illustrée à la figure 11.5 que l'on va étirer dans chaque direction. Les dimensions de la plaque sont les suivantes :  $-5 \leq x \leq 5$  et  $-2 \leq y \leq 2$  et le trou est de rayon 1 et centré en  $(0, 0)$ . On fait l'hypothèse de déformation plane. Sur le côté gauche de la plaque ( $x = -5$ ), on impose un déplacement de la forme  $\mathbf{u} = (-0,1, 0)$  tandis que de l'autre côté ( $x = 5$ ),

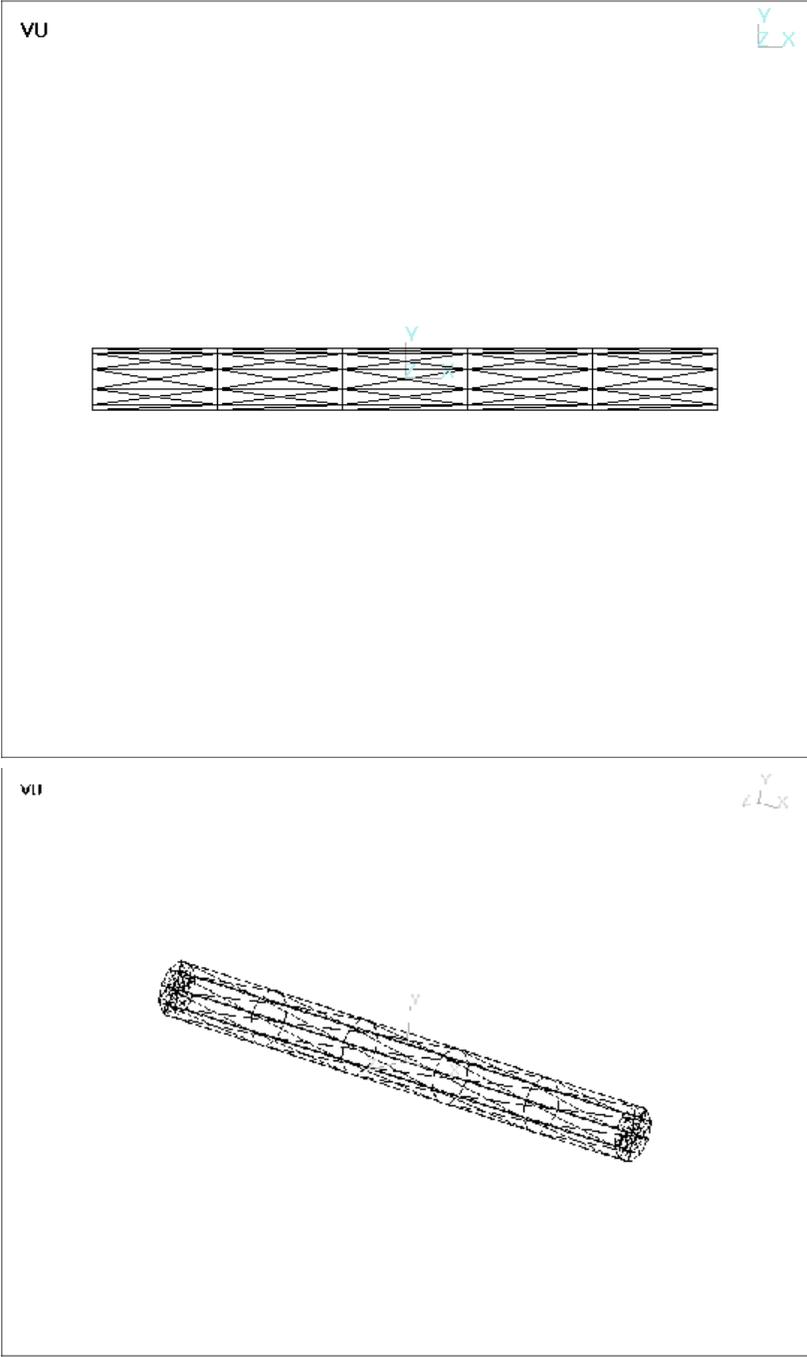


Figure 11.2 – Maillage utilisé

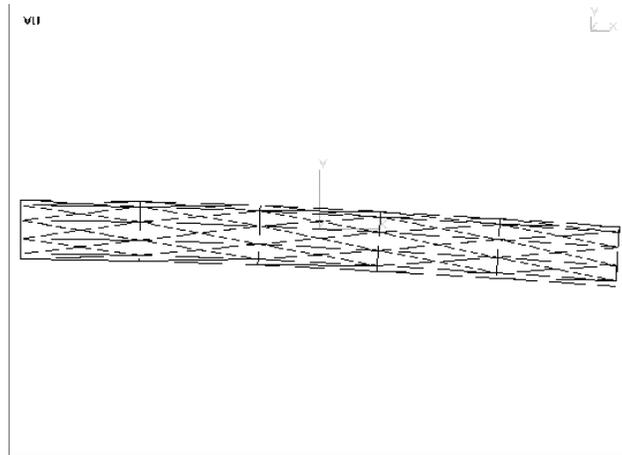


Figure 11.3 – Déformation du maillage

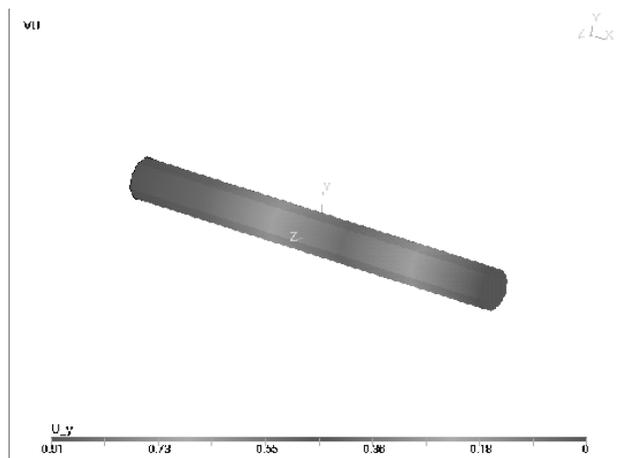


Figure 11.4 – Déplacement vertical

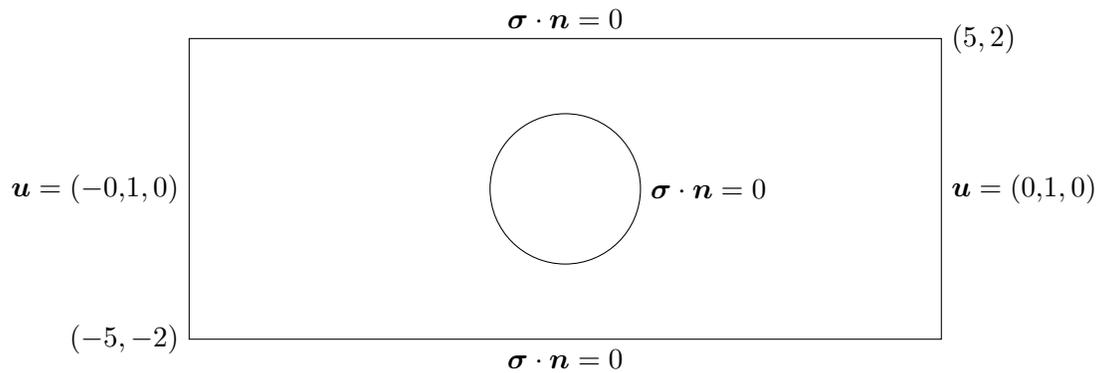


Figure 11.5 – Plaque trouée : géométrie et conditions aux limites

on impose  $\mathbf{u} = (0, 1, 0)$ . Aucune traction ( $\boldsymbol{\sigma} \cdot \mathbf{n} = 0$ ) n'est imposée sur les parois supérieures et inférieures, de même que sur la paroi du trou (le cercle).

Pour effectuer les simulations nous avons utilisé un élément quadrangulaire à 9 noeuds illustré à la figure 6.4. De manière à bien illustrer la faiblesse de la formulation en déplacement seulement (par rapport à la formulation mixte du chapitre 13), nous présentons aux figures 11.6 et 11.7 les résultats obtenus lorsque le coefficient de Poisson  $\nu$  tend vers  $1/2$ . On constate aisément que les contraintes deviennent très imprécises à  $\nu = 0,4999$ . Notons au passage que les déplacements de la figure 11.6 ont été amplifiés d'un facteur 3.

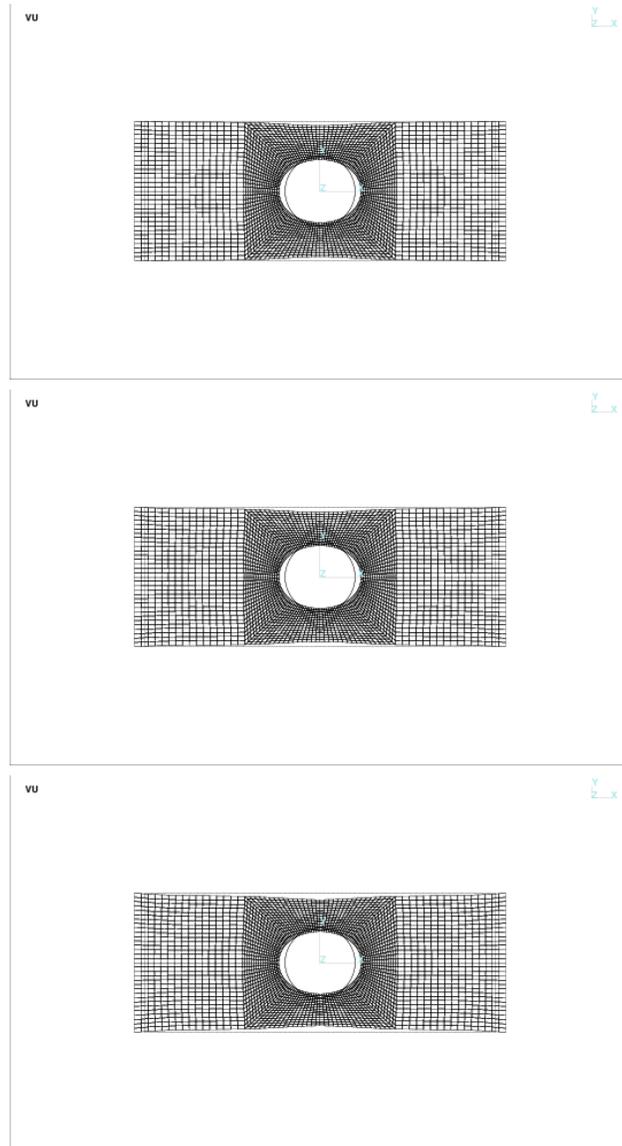


Figure 11.6 – Déplacements :  $\nu = 0,3$ ,  $\nu = 0,4$  et  $\nu = 0,4999$

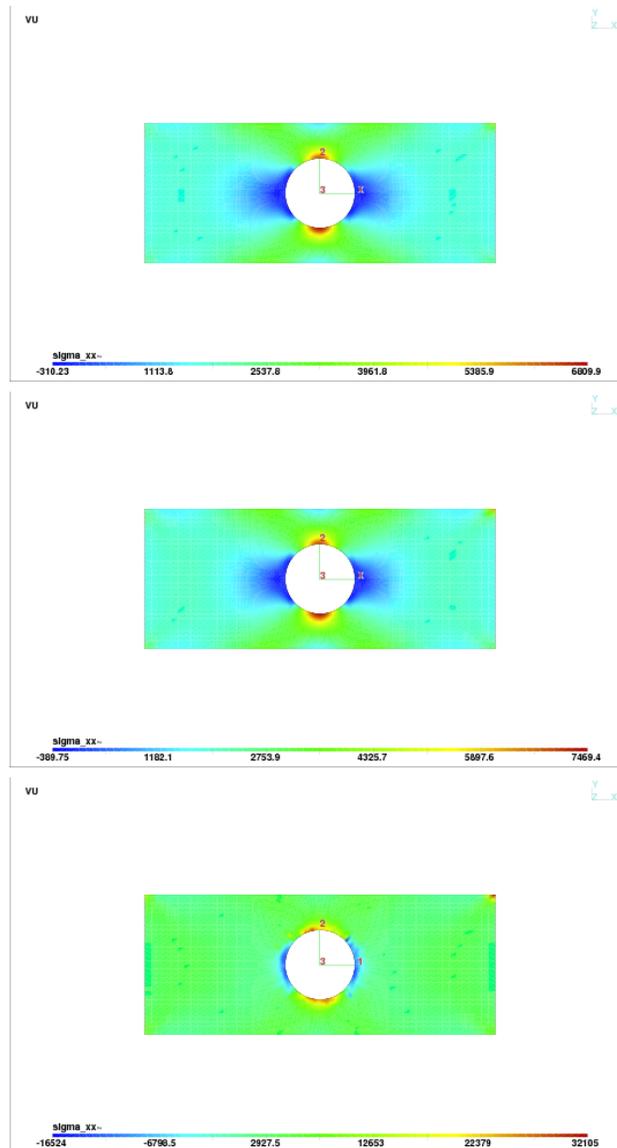


Figure 11.7 – Contrainte  $\sigma_{11}$  :  $\nu = 0,3$ ,  $\nu = 0,4$  et  $\nu = 0,4999$

### 11.14 Exercices

1. Montrer que  $\text{tr}(\boldsymbol{\gamma}(\mathbf{u})) = \nabla \cdot \mathbf{u}$ .
2. Montrer que si  $C_{ijkl} = \lambda I_{ij}I_{kl} + \mu(I_{ik}I_{jl} + I_{il}I_{jk})$ , alors

$$\mathbf{C} : \boldsymbol{\gamma}(\mathbf{u}) = \lambda(\nabla \cdot \mathbf{u})\mathbf{I} + 2\mu\boldsymbol{\gamma}(\mathbf{u})$$

3. Pour un problème d'élasticité linéaire en formulation mixte, la condition naturelle est  $\boldsymbol{\sigma} \cdot \mathbf{n}$  où :

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\gamma}(\mathbf{u}) - \frac{2\mu}{3}(\nabla \cdot \mathbf{u})\mathbf{I} - p\mathbf{I}$$

Si on suppose maintenant que nous avons une paroi verticale pour un problème bidimensionnel, donner l'expression de la condition naturelle  $\boldsymbol{\sigma} \cdot \mathbf{n}$  sur cette paroi.

# Chapitre 12

## Problème de Stokes

### 12.1 Introduction

Le problème de « Stokes » est au coeur de la simulation numérique en mécanique des fluides. Une bonne compréhension des difficultés liées à sa discrétisation par la méthode des éléments finis ouvre la voie à toute une panoplie d'applications. Cela va des écoulements de fluides fortement visqueux avec des applications notamment à la mise en forme des polymères, aux fluides peu visqueux que nous modéliserons à l'aide des équations de Navier-Stokes pour étudier des applications en hydraulique par exemple.

Introduisons en premier lieu un peu de notation. On note  $\boldsymbol{\sigma}$  le tenseur de contraintes de Cauchy que l'on peut par la suite écrire comme une somme de différentes contributions :

$$\boldsymbol{\sigma} = -p\mathbf{I} + \boldsymbol{\tau}$$

où  $p$  est la pression et  $\boldsymbol{\tau}$  est le tenseur des extra-contraintes. L'équation d'équilibre s'écrit encore ici :

$$-\nabla \cdot \boldsymbol{\sigma} = \mathbf{r}$$

où  $\mathbf{r}$  est un vecteur de forces volumiques supposé connu. En tout point du domaine de calcul, nous chercherons à déterminer le champ de vitesse  $\mathbf{u} = (u_1, u_2, u_3)$  de même que la pression  $p$ . Nous nous intéressons en particulier au cas des fluides incompressibles pour lesquels la conservation de la masse impose que :

$$\nabla \cdot \mathbf{u} = 0$$

Pour fermer le système, il reste à établir une relation entre le tenseur des extra-contraintes  $\boldsymbol{\tau}$  et le champ de vitesse  $\mathbf{u}$  ou plus précisément le tenseur de taux de déformation  $\dot{\boldsymbol{\gamma}}(\mathbf{u})$ . Cette relation, dite *loi de comportement*, dépend de la nature du fluide considéré. Le premier cas (et aussi le plus simple) concerne les fluides dits *newtoniens* pour lesquels la loi de comportement s'écrit :

$$\boldsymbol{\tau} = 2\eta\dot{\boldsymbol{\gamma}}(\mathbf{u}) \tag{12.1}$$

où  $\eta$  est la viscosité du fluide que l'on suppose constante. On note  $\dot{\gamma}(\mathbf{u})$  le tenseur de taux de déformation défini par :

$$\dot{\gamma}(\mathbf{u}) = \begin{bmatrix} \frac{\partial u_1}{\partial x} & \frac{1}{2} \left( \frac{\partial u_1}{\partial y} + \frac{\partial u_2}{\partial x} \right) & \frac{1}{2} \left( \frac{\partial u_1}{\partial z} + \frac{\partial u_3}{\partial x} \right) \\ \frac{1}{2} \left( \frac{\partial u_2}{\partial x} + \frac{\partial u_1}{\partial y} \right) & \frac{\partial u_2}{\partial y} & \frac{1}{2} \left( \frac{\partial u_2}{\partial z} + \frac{\partial u_3}{\partial y} \right) \\ \frac{1}{2} \left( \frac{\partial u_3}{\partial x} + \frac{\partial u_1}{\partial z} \right) & \frac{1}{2} \left( \frac{\partial u_3}{\partial y} + \frac{\partial u_2}{\partial z} \right) & \frac{\partial u_3}{\partial z} \end{bmatrix}$$

Notons la différence entre le tenseur de taux de déformation  $\dot{\gamma}(\mathbf{u})$  (portant sur un champ de vitesses) et le tenseur de déformation  $\gamma(\mathbf{u})$  du chapitre précédent qui portait sur un champ de déplacement. Le tenseur de déformation n'a pas d'unités tandis que celles du tenseur de taux de déformation sont des  $s^{-1}$ . L'écriture des 2 tenseurs est cependant identique.

Les fluides newtoniens sont donc caractérisés par une viscosité constante, ce qui convient dans certaines applications mais pas du tout dans d'autres. On vérifie en effet facilement que la viscosité de bon nombre de fluides n'est pas constante mais varie avec le cisaillement (et la température). Plus précisément, plus un fluide est cisailé (par exemple entre 2 plaques en mouvement l'une par rapport à l'autre), plus la viscosité diminue. On doit également ajouter une forte dépendance de la viscosité en fonction de la température. Les automobilistes savent bien que l'huile à moteur est beaucoup plus visqueuse lorsque le moteur est froid que lorsqu'il est chaud. Dans un premier temps, nous ne nous intéresserons qu'au cas où la viscosité est constante. Nous reviendrons sur un cas plus général à la section 12.8.

## 12.2 Le problème continu

La simulation numérique des procédés de mise en forme des matériaux nous amène à étudier de manière approfondie les méthodes d'éléments finis pour le problème de Stokes qui régit l'écoulement des fluides à forte viscosité. C'est bien sûr le cas des polymères. Rappelons la forme générale des équations de conservation de la quantité de mouvement et de la masse :

$$\begin{cases} -\nabla \cdot (\boldsymbol{\tau} - p\mathbf{I}) & = \mathbf{r} \\ \nabla \cdot \mathbf{u} & = 0 \end{cases} \quad (12.2)$$

De la relation 12.1, on a :

$$\begin{cases} -\nabla \cdot (2\eta\dot{\gamma}(\mathbf{u}) - p\mathbf{I}) & = \mathbf{r} \\ \nabla \cdot \mathbf{u} & = 0 \end{cases} \quad (12.3)$$

auquelles il faut ajouter des conditions aux limites appropriées. La résolution par la méthode des éléments finis nécessite l'obtention d'une formulation variationnelle du problème. On multiplie ces deux équations respectivement par des fonctions test  $\mathbf{w} \in V$  et  $q \in Q$  où  $V$  et  $Q$  sont des espaces fonctionnels appropriés qu'il n'est pas nécessaire de spécifier pour le moment. On intègre ensuite par parties sur le domaine  $\Omega$  de frontière  $\Gamma$  et on obtient ainsi  $\forall \mathbf{w} \in V$  et  $\forall q \in Q$  :

$$\int_{\Omega} (2\eta\dot{\gamma}(\mathbf{u}) : \dot{\gamma}(\mathbf{w}) - p\nabla \cdot \mathbf{w}) \, dv = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv + \int_{\Gamma} ((-p\mathbf{I} + 2\eta\dot{\gamma}(\mathbf{u})) \cdot \mathbf{n}) \cdot \mathbf{w} \, d\Gamma \tag{12.4}$$

$$\int_{\Omega} q\nabla \cdot \mathbf{u} \, dv = 0$$

Un coup d'oeil au terme de bord fait ressortir deux types de conditions aux limites. De façon naturelle, c'est-à-dire par intégration par parties, on a fait apparaître sur la frontière  $\Gamma$  l'expression  $(-p\mathbf{I} + 2\eta\dot{\gamma}(\mathbf{u})) \cdot \mathbf{n} = \boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{t}$  qui devient ainsi la *condition naturelle* du problème. Le terme  $\boldsymbol{\sigma} \cdot \mathbf{n}$  représente en fait le *vecteur de contrainte* que l'on assimile à une traction (force par unité de surface). L'imposition de la traction est en dualité avec la *condition essentielle* qui porte sur les composantes de  $\mathbf{u}$ . Les conditions aux limites que l'on peut imposer sont donc :

- toutes les composantes de vitesse :  $\mathbf{u} = \mathbf{u}_0$  pour  $\mathbf{u}_0$  donné ;
- toutes les composantes de la contrainte normale :  $\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{t}$  pour  $\mathbf{t}$  donné ;
- des conditions mixtes comme par exemple :  $u_1 = u_1^0$  de même que les deuxième et troisième composantes de la contrainte normale  $(\boldsymbol{\sigma} \cdot \mathbf{n})_2 = t_2$  et  $(\boldsymbol{\sigma} \cdot \mathbf{n})_3 = t_3$  ;
- toute autre combinaison des composantes de vitesse  $u_i$  et des composantes de la contrainte normale  $(\boldsymbol{\sigma} \cdot \mathbf{n})_j$ , pourvu que  $i$  soit différent de  $j$ .

Comme toujours, on n'impose jamais une condition essentielle ( $u_j$ ) et une condition naturelle ( $t_j$ ) sur une même partie de la frontière. Notons enfin qu'il n'y a pas de condition essentielle sur la pression  $p$ . On ne peut imposer une pression (ou un gradient de pression) que par le biais de la condition naturelle.

Sur la partie  $\Gamma_0$  de la frontière  $\Gamma$  où le champ de vitesse  $\mathbf{u}$  (ou l'une de ses composantes  $u_i$ ) est connu, on a une condition essentielle et la fonction test  $\mathbf{w}$  (ou sa composante  $w_i$ ) s'annule sur cette partie de la frontière. Par contre, sur le complément  $\Gamma_1$  de  $\Gamma_0$ , on doit connaître la condition naturelle  $(\boldsymbol{\sigma} \cdot \mathbf{n})$  et les fonctions test  $\mathbf{w}$  ne doivent pas toutes s'y annuler. On obtient ainsi :

$$\int_{\Omega} (2\eta\dot{\gamma}(\mathbf{u}) : \dot{\gamma}(\mathbf{w}) - p\nabla \cdot \mathbf{w}) \, dv = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv + \int_{\Gamma_1} \mathbf{t} \cdot \mathbf{w} \, d\Gamma \quad \forall \mathbf{w} \in V \tag{12.5}$$

$$\int_{\Omega} q\nabla \cdot \mathbf{u} \, dv = 0 \quad \forall q \in Q$$

Nous supposons par la suite des conditions essentielles homogènes ( $\mathbf{u} = \mathbf{w} = 0$ ) partout sur la frontière  $\Gamma$ , le cas inhomogène ne posant aucune difficulté supplémentaire puisqu'il est possible d'effectuer un relèvement de la condition essentielle. Il en résulte un problème plus simple mais dont la généralité nous permettra d'aborder facilement les cas plus complexes. Par exemple, le cas

des fluides viscoplastiques se traite de manière similaire, bien que dans ce cas, le problème soit non linéaire. Le problème consiste donc à trouver  $\mathbf{u} \in V$  et  $p \in Q$  ( $V$  et  $Q$  seront précisés plus loin) tels que :

$$\begin{cases} \int_{\Omega} 2\eta\dot{\gamma}(\mathbf{u}) : \dot{\gamma}(\mathbf{w}) \, dv - \int_{\Omega} p\nabla \cdot \mathbf{w} \, dv = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv & \forall \mathbf{w} \in V \\ \int_{\Omega} q\nabla \cdot \mathbf{u} \, dv = 0 & \forall q \in Q \end{cases} \quad (12.6)$$

On a introduit un signe négatif dans la deuxième équation pour des raisons de symétrie. On introduit maintenant les formes bilinéaires  $a$  et  $b$  définies respectivement sur  $V \times V$  et  $V \times Q$  de sorte qu'on peut écrire :

$$\begin{cases} a(\mathbf{u}, \mathbf{w}) + b(\mathbf{w}, p) = (\mathbf{r}, \mathbf{w}) & \forall \mathbf{w} \in V \\ b(\mathbf{u}, q) = 0 & \forall q \in Q \end{cases} \quad (12.7)$$

### 12.2.1 Cadre fonctionnel

L'analyse numérique du système 12.6 (ou 12.7) est relativement complexe. Tout d'abord, les espaces fonctionnels sont  $V = (H_0^1(\Omega))^d$  ( $d$  étant la dimension d'espace) et  $Q = L^2(\Omega)$ , l'ensemble des fonctions de carré sommable. La forme bilinéaire  $b$  définit implicitement deux applications linéaires. La première, que nous noterons  $B$ , est définie sur  $V$  à valeur dans  $Q'$ , le dual de  $Q$  :

$$\begin{aligned} B : V &\longrightarrow Q' \\ \mathbf{w} &\longrightarrow B\mathbf{w} \end{aligned}$$

En effet, il suffit de poser :

$$\langle B\mathbf{w}, q \rangle = b(\mathbf{w}, q) = - \int_{\Omega} q\nabla \cdot \mathbf{w} \, dv \quad \forall q \in Q$$

et on vérifie facilement (voir les exercices de fin de chapitre) qu'il s'agit bien d'une forme linéaire continue sur  $Q$ . On peut aussi définir le noyau de  $B$  (noté  $\ker B$ ) par :

$$\begin{aligned} \ker B &= \{\mathbf{w} \in V \mid B\mathbf{w} = 0\} \\ &= \{\mathbf{w} \in V \mid \langle B\mathbf{w}, q \rangle = b(\mathbf{w}, q) = 0 \quad \forall q \in Q\} \\ &= \left\{ \mathbf{w} \in (H_0^1(\Omega))^d \mid - \int_{\Omega} q\nabla \cdot \mathbf{w} \, dv = 0 \quad \forall q \in Q \right\} \\ &= \left\{ \mathbf{w} \in (H_0^1(\Omega))^d \mid \nabla \cdot \mathbf{w} = 0 \right\} \end{aligned}$$

#### ◆ Remarque 12.1

Si  $\mathbf{u} \in (H_0^1(\Omega))^d$ , on remarque que  $\nabla \cdot \mathbf{u} \in L^2(\Omega)$  et que la deuxième équation du système 12.6 est parfaitement équivalente à imposer  $\nabla \cdot \mathbf{u} = 0$ . ♦

On peut définir une deuxième application linéaire, cette fois de  $Q$  à valeur dans  $V'$ , dual de  $V$ , que nous noterons  $B^T$  et qui est définie par :

$$\begin{aligned} B^T : Q &\longrightarrow V' \\ q &\longrightarrow B^T q \end{aligned}$$

Comme précédemment, on pose :

$$\langle \mathbf{w}, B^T q \rangle = b(\mathbf{w}, q) = - \int_{\Omega} q \nabla \cdot \mathbf{w} \, dv \quad \forall \mathbf{w} \in V$$

Il est alors évident que :

$$\langle B\mathbf{w}, q \rangle = \langle \mathbf{w}, B^T q \rangle \quad \forall \mathbf{w} \in V, \forall q \in Q$$

De même, on pose :

$$\begin{aligned} \ker B^T &= \{q \in Q \mid B^T q = 0\} \\ &= \{q \in Q \mid \langle \mathbf{w}, B^T q \rangle = b(\mathbf{w}, q) = 0 \quad \forall \mathbf{w} \in V\} \\ &= \left\{ q \in Q \mid - \int_{\Omega} q \nabla \cdot \mathbf{w} \, dv = 0 \quad \forall \mathbf{w} \in (H_0^1(\Omega))^d \right\} \end{aligned}$$

En intégrant par parties (puisque les fonctions de  $(H_0^1(\Omega))^d$  s'annulent au bord), on a :

$$\ker B^T = \{q \in Q \mid \int_{\Omega} \nabla q \cdot \mathbf{w} \, dv = 0 \quad \forall \mathbf{w} \in (H_0^1(\Omega))^d\}$$

ce qui entraîne que  $\nabla q = 0$  et donc que  $\ker B^T$  est constitué des fonctions constantes sur  $\Omega$ .

### 12.2.2 Existence et unicité

Les théorèmes qui suivent, et que nous ne démontrerons pas, sont d'une grande importance théorique.

**Théorème 1** (Existence et unicité du problème continu). *Sous les conditions suivantes :*

- Les formes bilinéaires  $a$  et  $b$  sont continues respectivement sur  $V \times V$  et  $V \times Q$  ;
- La forme bilinéaire  $a$  est coercive sur  $\ker B$  c.-à-d. :

$$a(\mathbf{w}, \mathbf{w}) \geq \alpha \|\mathbf{w}\|_V^2 \quad \forall \mathbf{w} \in \ker B \tag{12.8}$$

– Il existe une constante positive  $\beta$  telle que la forme bilinéaire  $b$  vérifie :

$$\inf_{q \in Q} \sup_{\mathbf{w} \in V} \frac{b(\mathbf{w}, q)}{\|q\|_Q \|\mathbf{w}\|_V} \geq \beta \quad (12.9)$$

alors la solution  $(\mathbf{u}, p)$  du problème de Stokes 12.6 (ou 12.7) est unique dans  $V \times Q / \ker B^T$ .

**Démonstration :** Voir Brezzi-Fortin [7].

La condition de coercivité 12.8 découle encore une fois de l'inégalité de Korn 11.15. La forme bilinéaire  $a$  est même coercive sur tout l'espace  $V$  et pas seulement sur le noyau. On peut aussi montrer que l'inégalité 12.9 est bien vérifiée. On remarque cependant que la pression n'est déterminée qu'à une fonction de  $\ker B^T$  près c'est-à-dire à une constante près, d'où la notation  $Q / \ker B^T$ . Pour s'en convaincre, il suffit de revenir au problème initial 12.2 et de remarquer qu'on peut remplacer  $p$  par  $p + c$  sans rien changer. Ceci n'est par contre vrai que dans le cas de conditions essentielles (sur  $\mathbf{u}$ ) imposées partout sur la frontière. En effet, si on impose une condition naturelle, la pression  $p$  apparaissant explicitement dans  $\boldsymbol{\sigma} \cdot \mathbf{n}$ , elle n'est plus déterminée à une constante près.

### 12.3 Le problème discret

La discrétisation par éléments finis du problème de Stokes suit les étapes habituelles. On approxime  $(\mathbf{u}, p)$  par des fonctions  $(\mathbf{u}_h, p_h) \in V_h \times Q_h$ . Les sous-espaces  $V_h$  de  $V$  et  $Q_h$  de  $Q$  sont de dimension finie et seront déterminés plus précisément plus loin. Le problème discrétisé s'écrit alors :

$$\begin{aligned} a(\mathbf{u}_h, \mathbf{w}_h) + b(\mathbf{w}_h, p_h) &= (\mathbf{r}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in V_h \\ b(\mathbf{u}_h, q_h) &= 0 \quad \forall q_h \in Q_h \end{aligned} \quad (12.10)$$

La condition d'incompressibilité discrétisée s'écrit maintenant :

$$-\int_{\Omega} q_h \nabla \cdot \mathbf{w}_h \, dv = 0 \quad \forall q_h \in Q_h$$

et n'est pas équivalente à  $\nabla \cdot \mathbf{w}_h = 0$ . Pour que l'on puisse conclure que  $\nabla \cdot \mathbf{w}_h = 0$ , il faudrait que  $q_h$  parcourt tout l'espace  $Q$  et pas seulement un sous-espace  $Q_h$ . Comme dans le cas continu, on peut définir :

$$\ker B_h = \left\{ \mathbf{w}_h \in V_h \mid -\int_{\Omega} q_h \nabla \cdot \mathbf{w}_h \, dv = 0 \quad \forall q_h \in Q_h \right\}$$

De même, si on pose :

$$\ker B_h^T = \left\{ q_h \in Q_h \mid -\int_{\Omega} q_h \nabla \cdot \mathbf{w}_h \, dv = 0 \quad \forall \mathbf{w}_h \in V_h \right\}$$

on peut construire, pour certains choix d'espaces  $V_h$  et  $Q_h$  des fonctions de  $\ker B_h^T$  qui ne sont pas constantes sur  $\Omega$  et qui peuvent altérer la solution numérique (voir les exercices de fin de chapitre). C'est le cas par exemple des pressions dites en damier (« checkerboard pressure » en anglais) que l'on retrouve dans certains cas. Il en résulte que le choix des espaces  $V_h$  et  $Q_h$  est assez délicat.

### 12.3.1 Quelques résultats théoriques importants

Les résultats qui suivent, que nous ne démontrerons pas, ont une grande importance théorique et pratique. Nous nous efforcerons d'illustrer ces résultats et d'en approfondir la compréhension au moyen de quelques exemples simples.

**Théorème 2** (Existence et unicité du problème discret). *Sous les conditions suivantes :*

– la forme bilinéaire  $a$  est coercive sur  $\ker B_h$  :

$$a(\mathbf{w}_h, \mathbf{w}_h) \geq \alpha \|\mathbf{w}_h\|_V^2 \quad \forall \mathbf{w}_h \in \ker B_h$$

– la forme bilinéaire  $b$  vérifie :

$$\inf_{q_h \in Q_h} \sup_{\mathbf{w}_h \in V_h} \frac{b(\mathbf{w}_h, q_h)}{\|q_h\|_Q \|\mathbf{w}_h\|_V} \geq \beta_h \tag{12.11}$$

alors la solution  $(\mathbf{u}_h, p_h)$  du problème de Stokes discret 12.10 est unique dans  $V_h \times Q_h / \ker B_h^T$ .

**Démonstration :** Voir Brezzi-Fortin [7].

Ici encore, on peut montrer que les 2 conditions sont bien vérifiées et ce, quel que soit le choix des espaces de discrétisation  $V_h$  et  $Q_h$ . La condition 12.11 est en fait équivalente à affirmer que l'image de  $B_h$  est fermée, ce qui est toujours le cas en dimension finie. Il est aussi primordial que  $\ker B_h^T \subset \ker B^T$  car autrement, la pression ne sera plus déterminée à une constante près et risque d'être perturbée par la présence des éléments supplémentaires contenus dans le noyau de  $B_h^T$ .

Un autre problème provient de la constante  $\beta_h$  présente dans la condition 12.11 et qui peut dépendre de  $h$ . Cette constante se retrouve, comme nous le verrons à la prochaine section, dans la borne d'erreur de sorte que l'on peut perdre la convergence de la solution discrète vers la solution continue. Pour éviter ce problème, on doit montrer que

$$\inf_{q_h \in Q_h} \sup_{\mathbf{w}_h \in V_h} \frac{b(\mathbf{w}_h, q_h)}{\|q_h\|_{0,\Omega} \|\mathbf{w}_h\|_{1,\Omega}} \geq \beta_0 > 0 \tag{12.12}$$

où la constante  $\beta_0$  est cette fois indépendante de  $h$ . C'est la condition de Brezzi-Babüska dite parfois condition *inf-sup*.

**Théorème 3** (Convergence). *Sous les hypothèses d'existence et d'unicité des solutions  $(\mathbf{u}, p)$  et  $(\mathbf{u}_h, p_h)$  des problèmes de Stokes continu et discret et si la condition de Brezzi-Babüska 12.12 est vérifiée, alors :*

$$\|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega} + \|p - p_h\|_{0,\Omega} \leq C \left[ \inf_{\mathbf{w}_h \in V_h} \|\mathbf{u} - \mathbf{w}_h\|_{1,\Omega} + \inf_{q_h \in Q_h} \|p - q_h\|_{0,\Omega} \right] \tag{12.13}$$

◆ **Remarque 12.2**

La constante  $C$  apparaissant dans le théorème précédent dépend entre autres choses de  $1/\alpha$  et de  $1/\beta_0$  d'où l'importance que cette dernière constante ne dépende pas de  $h$ . C'est pourquoi, bien que

la condition 12.11 soit toujours vérifiée, on n'a pas toujours convergence vers la solution exacte et le théorème 12.13 n'est valide que si la condition *inf-sup* 12.12 est vraie pour  $\beta_0$  indépendant de  $h$ .

◆

**Théorème 4** (Ordre de convergence). *Sous les mêmes hypothèses qu'au théorème précédent et si de plus le sous-espace  $V_h$  contient les polynômes de degré  $k$  et le sous-espace  $Q_h$  contient les polynômes de degré  $(k - 1)$ , alors :*

$$\| \mathbf{u} - \mathbf{u}_h \|_{1,\Omega} + \| p - p_h \|_{0,\Omega} \leq Ch^k (\| \mathbf{u} \|_{k+1,\Omega} + \| p \|_{k,\Omega}) \quad (12.14)$$

◆ **Remarque 12.3**

Il est important de remarquer qu'il ne suffit pas de prendre des polynômes de degré  $k$  pour la vitesse et de degré  $k - 1$  pour la pression pour avoir une discrétisation convergente, comme cela est parfois présenté dans la littérature. La condition Brezzi-Babůska 12.12 doit aussi être vérifiée. ◆

◆ **Remarque 12.4**

La transformation géométrique  $F^K$  de l'élément de référence  $\hat{K}$  vers  $K$  joue un rôle important pour déterminer si un espace  $V_h$  contient ou non les polynômes de degré  $k$ . ... ◆

## 12.4 Formulation point-selle (facultative)

Cette section présente comment on peut interpréter le système de Stokes comme un problème de point-selle. La lecture de cette section est facultative mais permettra au lecteur une compréhension encore plus profonde et ouvre la porte à la théorie des points-selles et aux algorithmes qui s'y rattachent.

Une autre façon d'aborder le problème de Stokes consiste à travailler directement dans le noyau de  $B$  c'est-à-dire avec les fonctions à divergence nulle. Revenant au système 12.3, on multiplie par une fonction  $\mathbf{w} \in \ker B$  et on intègre par parties. Le problème de Stokes peut maintenant s'écrire :

♠ Trouver  $\mathbf{u} \in \ker B$  tel que :

$$a(\mathbf{u}, \mathbf{w}) = (\mathbf{r}, \mathbf{w}) \quad \forall \mathbf{w} \in \ker B \quad (12.15)$$

Classiquement (voir le théorème 3.5), ce problème est équivalent à :

$$\inf_{\mathbf{w} \in \ker B} \frac{1}{2} a(\mathbf{w}, \mathbf{w}) - (\mathbf{r}, \mathbf{w})$$

Le théorème de Lax-Milgram assure l'existence et l'unicité de la solution  $\mathbf{u}$  sous l'hypothèse habituelle de coercivité de la forme bilinéaire  $a$  sur  $\ker B$  (et non sur tout l'espace  $V$ ). Par contre, démontrer l'existence et l'unicité de la pression  $p$  est plus délicat. Nous y reviendrons plus loin.

On peut de toute évidence pénaliser la contrainte d'incompressibilité et il en résulte un problème de point-selle du *lagrangien*  $\mathcal{L}$  défini par :

$$\mathcal{L}(\mathbf{w}, q) = \frac{1}{2}a(\mathbf{w}, \mathbf{w}) + b(\mathbf{w}, q) - (\mathbf{r}, \mathbf{w})$$

En dérivant (au sens de Gâteaux et de l'équation 8.10) par rapport à chacune des variables, on obtient aisément la formulation variationnelle 12.6. En effet, on caractérise le point-selle  $(\mathbf{u}, p)$  par les 2 relations :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \mathbf{u}}(\mathbf{u}, p) \cdot \mathbf{w} &= \frac{d}{d\lambda}(\mathcal{L}(\mathbf{u} + \lambda \mathbf{w}, p)) \Big|_{\lambda=0} = 0 \\ \frac{\partial \mathcal{L}}{\partial p}(\mathbf{u}, p) \cdot q &= \frac{d}{d\lambda}(\mathcal{L}(\mathbf{u}, p + \lambda q)) \Big|_{\lambda=0} = 0 \end{aligned}$$

qui ne sont rien d'autre que la formulation 12.6. Le problème de Stokes est donc parfaitement équivalent à la recherche d'un point-selle :

$$\inf_{\mathbf{w} \in V} \sup_{q \in Q} \mathcal{L}(\mathbf{w}, q)$$

On peut aussi introduire le *lagrangien augmenté* noté  $\mathcal{L}_r$  et qui est défini par :

$$\mathcal{L}_r(\mathbf{w}, q) = \mathcal{L}(\mathbf{w}, q) + \frac{r}{2}(\nabla \cdot \mathbf{w}, \nabla \cdot \mathbf{w})$$

La valeur du paramètre  $r$  sera souvent prise très grande de manière à forcer la condition d'incompressibilité. Le point-selle du lagrangien augmenté est caractérisé par :

$$\begin{aligned} a(\mathbf{u}, \mathbf{w}) + b(\mathbf{w}, p) + r(\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{w}) &= (\mathbf{r}, \mathbf{w}) \quad \forall \mathbf{w} \in V \\ b(\mathbf{u}, q) &= 0 \quad \forall q \in Q \end{aligned} \tag{12.16}$$

et puisqu'au point-selle  $(\mathbf{u}, p)$ , on a  $\nabla \cdot \mathbf{u} = 0$ , la *formulation variationnelle pénalisée* 12.16 est équivalente au problème de Stokes 12.7.

## 12.5 Formulation variationnelle élémentaire

Nous suivrons une démarche similaire à celle du chapitre 11. On déduit facilement la formulation variationnelle élémentaire :

$$\begin{aligned} \int_K 2\mu \dot{\boldsymbol{\gamma}}(\mathbf{u}) : \dot{\boldsymbol{\gamma}}(\mathbf{w}) \, dv - \int_K p \nabla \cdot \mathbf{w} \, dv &= \int_K \mathbf{r}^K \cdot \mathbf{w} \, dv + \int_{\partial K} \mathbf{t}^K \cdot \mathbf{w} \, ds \\ \int_K q \nabla \cdot \mathbf{u} \, dv &= 0 \end{aligned} \tag{12.17}$$

Les approximations des vitesses et de la pression pouvant être de degrés différents, on pose :

$$\mathbf{u}(\mathbf{x})|_K \simeq \mathbf{u}_h(\mathbf{x})|_K = \mathbf{u}^K(\mathbf{x}) = \sum_{j=1}^{n_D^u} \alpha_j^K \Psi_{u,j}^K \quad p(\mathbf{x})|_K \simeq p_h(\mathbf{x})|_K = p^K(\mathbf{x}) = \sum_{j=1}^{n_D^p} p_j^K \psi_{p,j}^K$$

Le vecteur  $\alpha^K$  est défini au chapitre 11 à l'équation 11.26. Les  $n_D^u$  fonctions de base (vectorielles) en vitesse  $\Psi_{u,j}^K$  et les  $n_D^p$  fonctions de base en pression peuvent être de nature différente. Par exemple, les  $\Psi_{u,j}^K$  pourront être quadratiques et les  $\psi_{p,j}^K$  linéaires sur un élément. Il en résulte que le nombre de degrés de liberté pourra varier d'un noeud de calcul à l'autre et que les composantes de la vitesse et la pression pourront être évaluées en des noeuds de calcul différents. Le choix particulier des espaces de discrétisation sera précisé plus loin.

On peut de plus décomposer le système sous la forme :

$$\begin{bmatrix} A_{11}^K & A_{12}^K & A_{13}^K & (B_1^K)^T \\ A_{21}^K & A_{22}^K & A_{23}^K & (B_2^K)^T \\ A_{31}^K & A_{32}^K & A_{33}^K & (B_3^K)^T \\ B_1^K & B_2^K & B_3^K & 0 \end{bmatrix} \begin{bmatrix} U_1^K \\ U_2^K \\ U_3^K \\ P^K \end{bmatrix} = \begin{bmatrix} F_1^K \\ F_2^K \\ F_3^K \\ 0 \end{bmatrix} + \begin{bmatrix} S_1^K \\ S_2^K \\ S_3^K \\ 0 \end{bmatrix}$$

Les matrices  $A_{ij}^K$  sont de dimension  $n_C^u$  par  $n_C^u$ , où  $n_C^u$  est le nombre de noeuds de calcul en vitesse. On aura souvent  $n_D^u = d n_C^u$  où  $d$  est la dimension d'espace ( $d = 2$  ou  $3$ ). Les matrices  $B_i^K$  sont quant à elles de dimension  $n_D^p$  par  $n_C^u$  et sont donc rectangulaires dans le cas général.

On définit les fonctions de base en vitesse comme au chapitre 11 par l'équation 11.27. Les fonctions de base en pression sont obtenues de la manière habituelle. Les différentes matrices élémentaires prennent la forme :

$$(A_{11}^K)_{ij} = \int_K 2\mu \left( \frac{\partial \psi_{u,j}^K}{\partial x} \frac{\partial \psi_{u,i}^K}{\partial x} + \frac{1}{2} \frac{\partial \psi_{u,j}^K}{\partial y} \frac{\partial \psi_{u,i}^K}{\partial y} + \frac{1}{2} \frac{\partial \psi_{u,j}^K}{\partial z} \frac{\partial \psi_{u,i}^K}{\partial z} \right) dv$$

De même :

$$(A_{12}^K)_{ij} = \int_K \mu \left( \frac{\partial \psi_{u,j}^K}{\partial x} \frac{\partial \psi_{u,i}^K}{\partial y} \right) dv = (A_{21}^K)_{ji}$$

$$(A_{13}^K)_{ij} = \int_K \mu \left( \frac{\partial \psi_{u,j}^K}{\partial x} \frac{\partial \psi_{u,i}^K}{\partial z} \right) dv = (A_{31}^K)_{ji}$$

$$(A_{22}^K)_{ij} = \int_K 2\mu \left( \frac{1}{2} \frac{\partial \psi_{u,j}^K}{\partial x} \frac{\partial \psi_{u,i}^K}{\partial x} + \frac{\partial \psi_{u,j}^K}{\partial y} \frac{\partial \psi_{u,i}^K}{\partial y} + \frac{1}{2} \frac{\partial \psi_{u,j}^K}{\partial z} \frac{\partial \psi_{u,i}^K}{\partial z} \right) dv$$

$$(A_{23}^K)_{ij} = \int_K \mu \left( \frac{\partial \psi_{u,j}^K}{\partial y} \frac{\partial \psi_{u,i}^K}{\partial z} \right) dv = (A_{32}^K)_{ji}$$

$$(A_{33}^K)_{ij} = \int_K 2\mu \left( \frac{1}{2} \frac{\partial \psi_{u,j}^K}{\partial x} \frac{\partial \psi_{u,i}^K}{\partial x} + \frac{1}{2} \frac{\partial \psi_{u,j}^K}{\partial y} \frac{\partial \psi_{u,i}^K}{\partial y} + \frac{\partial \psi_{u,j}^K}{\partial z} \frac{\partial \psi_{u,i}^K}{\partial z} \right) dv$$

$$(B_1^K)_{ij} = \int_K \psi_{p,j}^K \frac{\partial \psi_{u,i}^K}{\partial x} dv$$

$$(B_2^K)_{ij} = \int_K \psi_{p,j}^K \frac{\partial \psi_{u,i}^K}{\partial y} dv$$

$$(B_3^K)_{ij} = \int_K \psi_{p,j}^K \frac{\partial \psi_{u,i}^K}{\partial z} dv$$

et enfin :

$$(f_1^K)_i = \int_K r_1^K \psi_{u,i}^K dv, \quad (S_1^K)_i = \int_{\partial K} t_1^K \psi_{u,i}^K ds$$

$$(f_2^K)_i = \int_K r_2^K \psi_{u,i}^K dv, \quad (S_2^K)_i = \int_{\partial K} t_2^K \psi_{u,i}^K ds$$

$$(f_3^K)_i = \int_K r_3^K \psi_{u,i}^K dv, \quad (S_3^K)_i = \int_{\partial K} t_3^K \psi_{u,i}^K ds$$

Il va se soit que ces intégrales sont évaluées sur l'élément de référence par le changement de variables approprié.

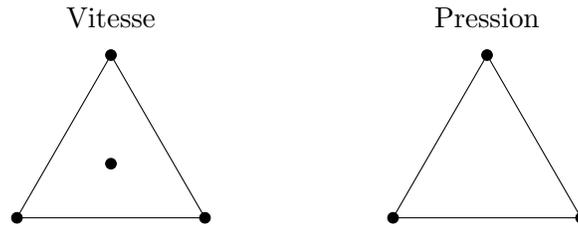
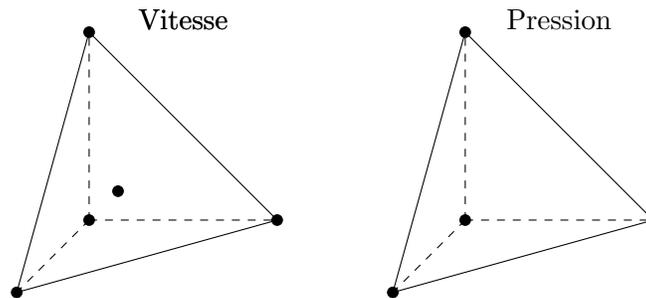
## 12.6 Choix des éléments

Nous présentons dans cette section des espaces discrets vérifiant la condition *inf-sup*. Les approximations en vitesse et pression vérifiant cette condition sont dites *compatibles*. Il existe plusieurs choix possibles et nous les classerons en deux catégories suivant que la discrétisation en pression est continue ou non. En effet, la pression  $p$  étant dans  $L^2(\Omega)$ , il n'est nullement nécessaire que sa discrétisation soit continue à la frontière des éléments.

### 12.6.1 Pression continue

La présentation la plus simple est graphique. On trouvera donc dans les figures qui suivent des exemples de discrétisations compatibles ainsi que leur ordre de convergence. On retrouvera des éléments triangulaires et quadrangulaires en dimension 2, de même que des éléments tétraédriques et hexaédriques en dimension 3.

Le premier élément compatible, illustré aux figures 12.1 et 12.2, est l'élément Mini décrit dans Arnold-Brezzi-Fortin [2]. La discrétisation de la pression est linéaire tandis que la vitesse est linéaire mais enrichie d'un noeud de calcul au centre de l'élément. Chaque élément contient donc 8 degrés de

Figure 12.1 – Élément Mini ( $O(h)$ )Figure 12.2 – Élément Mini en dimension 3 ( $O(h)$ )

liberté en vitesse et 3 en pression dans le cas bidimensionnel (respectivement 15 et 4 en dimension 3). Cet élément est très utilisé, particulièrement en dimension 3 (voir la figure 12.2) en raison de son faible nombre de degrés de liberté. Cet élément converge à l'ordre 1 car l'approximation en vitesse ne contient que les polynômes de degré 1.

Notons toutefois qu'il s'agit d'un élément peu précis. Il est en quelque sorte minimal en ce sens qu'il ne contient que tout juste ce qu'il faut pour être compatible.

Le deuxième élément illustré est celui de Taylor-Hood souvent appelé  $P_2 - P_1$  en raison de l'approximation quadratique des composantes de vitesse et de l'approximation  $P_1$  de la pression (voir les figures 12.3 et 12.4. Les noeuds de calcul sont situés aux sommets (vitesses et pression) et aux milieux des arêtes (vitesses seulement). Cet élément converge à l'ordre 2 ( $O(h^2)$ ) en vertu du théorème 12.14 car l'approximation en vitesse contient les polynômes de degré 2 et l'approximation en pression contient ceux d'ordre 1. Des équivalents quadrangulaires et hexaédriques, également d'ordre 2, sont illustrés aux figures 12.5 et 12.6.

#### ◆ Remarque 12.5

Dans le cas d'éléments à pression continue, il est parfois nécessaire d'imposer une condition supplémentaire sur le maillage pour avoir l'existence et l'unicité de la solution. On devra par exemple supposer qu'aucun élément n'a deux côtés sur la frontière du domaine (voir les exercices de fin de chapitre). ◆

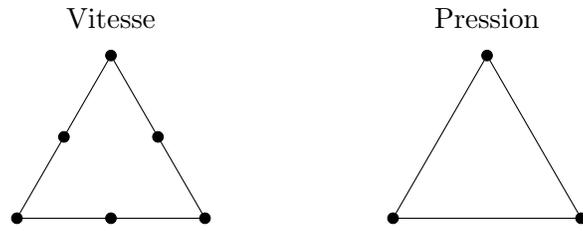


Figure 12.3 – Élément de Taylor-Hood  $P_2 - P_1$  ( $O(h^2)$ )

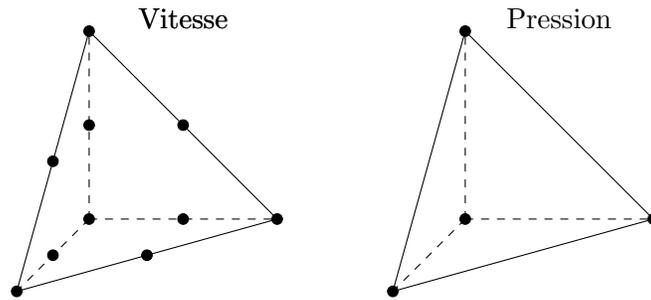


Figure 12.4 – Élément  $P_2 - P_1$  en dimension 3 ( $O(h^2)$ )

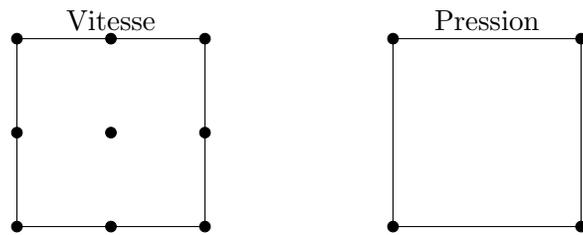
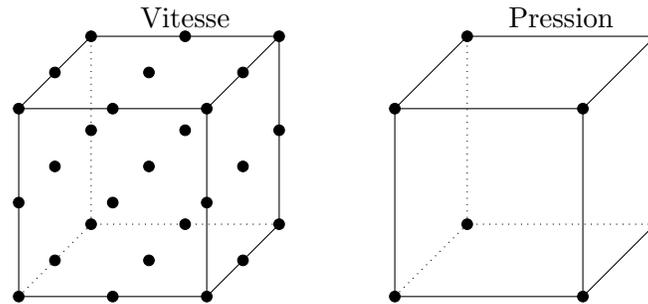
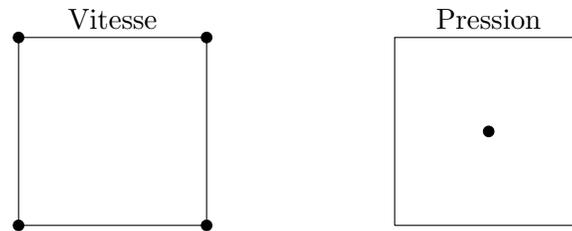


Figure 12.5 – Élément  $Q_2 - Q_1$  ( $O(h^2)$ )

Figure 12.6 – Élément  $Q_2 - Q_1$  en dimension 3 ( $O(h^2)$ )Figure 12.7 – Élément incompatible à pression discontinue  $Q_1 - Q_0$ 

### 12.6.2 Pression discontinue

Un des éléments les plus utilisés est le  $Q_1 - Q_0$  où les vitesses sont bilinéaires et la pression constante par élément. Cet élément est illustré à la figure 12.7 en dimension 2 et à la figure 12.8 en dimension 3. Malheureusement, cet élément est incompatible et nous en verrons les conséquences dans les exemples numériques.

Le premier élément compatible de la catégorie des éléments à pression discontinue est le  $P_2 - P_0$  présenté à la figure 12.9. Comme son nom l'indique, les composantes de vitesses sont quadratiques tandis que la pression est constante par élément. Le théorème 12.14 assure une convergence linéaire seulement car l'approximation en pression ne contient que les polynômes de degré 0. Il s'agit en fait d'un élément assez pauvre qui n'est pas beaucoup utilisé. C'est cependant l'un des plus simples vérifiant la condition *inf-sup*. Notons cependant que cet élément n'est valide qu'en dimension 2. En dimension 3, il faut mettre des degrés de liberté en vitesse sur le milieu des faces (nul besoin d'en mettre sur les arêtes) pour obtenir un élément compatible similaire.

L'élément de cette catégorie le plus connu est l'élément dit de Crouzeix et Raviart illustré à la figure 12.10 dans le cas bidimensionnel. En dimension 2, on approche les composantes de la vitesse

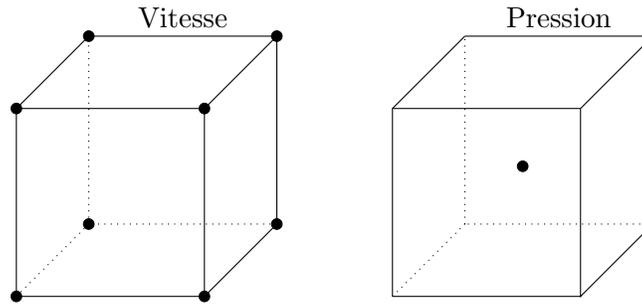


Figure 12.8 – Élément incompatible  $Q_1 - Q_0$  en dimension 3

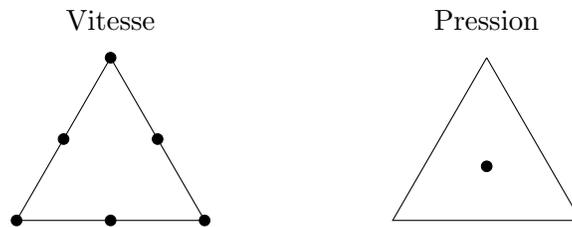


Figure 12.9 – Élément à pression discontinue  $P_2 - P_0$  ( $O(h)$ )

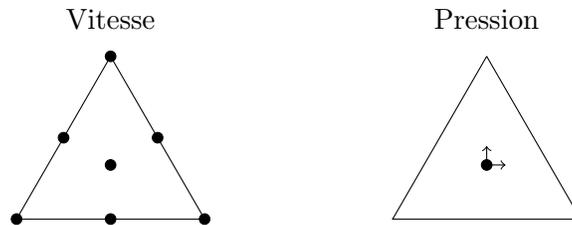


Figure 12.10 – Élément à pression discontinue de Crouzeix-Raviart  $P_2^+ - P_1$  ( $O(h^2)$ )

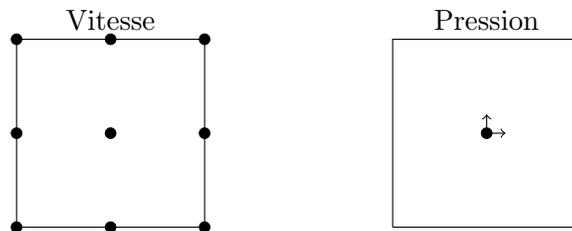


Figure 12.11 – Élément à pression discontinue  $Q_2 - P_1$  ( $O(h^2)$ )

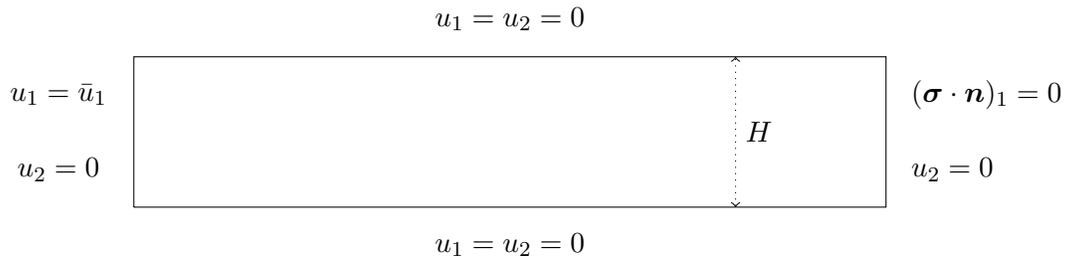


Figure 12.12 – Écoulement de Poiseuille (géométrie et conditions aux limites)

par des polynômes de degré 2 mais enrichi d’une fonction bulle de degré 3 associée à un noeud de calcul situé au centre de l’élément. Il y a donc 14 degrés de liberté en vitesse. La pression contient 3 degrés de liberté tous situés par exemple au barycentre de l’élément. Souvent, ces degrés de liberté sont associés à la pression et aux deux composantes du gradient de pression au barycentre. C’est ce que nous avons illustré par le symbole  $\blacklozenge$  à la figure 12.10.

En dimension 3, on ajoute aussi des fonctions bulles de degré 3 au centre de chaque face ainsi qu’une fonction bulle de degré 4 au barycentre de l’élément. On obtient ainsi 45 degrés de liberté en vitesse répartis sur 15 noeuds de calcul et 4 en pression (la pression et les 3 composantes du gradient au barycentre). Notons aussi les conséquences sur l’intégration numérique de la présence de ces fonctions bulles qui sont de degré élevé.

La pression est linéaire mais discontinue d’un élément à l’autre. C’est pourquoi nous avons illustré les noeuds de calcul en pression à l’intérieur de l’élément. Il en résulte que le nombre de degrés de liberté en pression est plus important que dans le cas d’une pression continue.

Notons enfin qu’il existe une technique, dite de condensation, qui permet d’éliminer localement les degrés de liberté en vitesse au barycentre de l’élément ainsi que le gradient de la pression, tout en conservant l’ordre de convergence de l’élément de Crouzeix et Raviart. En dimension 2, cette procédure réduit le système assemblé de  $4 \times Nel$  équations ( $6 \times Nel$  en dimension 3). On trouvera les détails dans Fortin et Fortin [19].

## 12.7 Résultats numériques : cas newtonien

### 12.7.1 L’écoulement de Poiseuille

Pour illustrer certains concepts que nous venons de présenter, il est utile de considérer un exemple très simple. Il s’agit de l’écoulement dit de Poiseuille dont la géométrie est illustrée à la figure 12.12. Ce problème décrit l’écoulement d’une fluide newtonien dans un canal dont les parois sont des plaques parallèles. Le fluide se déplace de gauche à droite dans une géométrie bidimensionnelle rectangulaire de longueur  $L$  ( $1 \leq x_1 \leq L$ ) et de hauteur  $H$  ( $0 \leq x_2 \leq H$ ) (les

effets tridimensionnels sont négligés). On impose en entrée (en  $x_1 = 0$ ) un profil parabolique :

$$\begin{aligned} u_1 &= \frac{6Q}{H^3} (x_2(H - x_2)) = \bar{u}_1 \\ u_2 &= 0 \end{aligned} \tag{12.18}$$

Dans l'expression de  $u_1$ ,  $Q$  désigne le débit imposé. On vérifie en effet que :

$$\int_0^H \bar{u}_1(x_2) dx_2 = Q$$

On impose en sortie (en  $x_1 = L$ ) que le profil est entièrement développé c.-à-d.  $u_2 = 0$  et que la première composante de la condition naturelle soit nulle  $(\boldsymbol{\sigma} \cdot \mathbf{n})_1 = 0$ . Sur les parois du canal (en  $x_2 = 0$  et  $x_2 = H$ ), une condition d'adhérence  $u_1 = u_2 = 0$  doit être imposée car un fluide visqueux adhère aux parois solides. La solution analytique de ce problème est tout simplement le déplacement de la condition aux limites imposée en entrée 12.18 dans tout le domaine. Quant à la pression, elle varie linéairement en fonction de  $x_1$  et s'écrit :

$$p = \left( \frac{12Q\eta}{H^3} \right) x_1 + C \tag{12.19}$$

Il suffit pour s'en convaincre de remplacer les relations 12.18 et 12.19 dans le système 12.3.

Nous avons résolu ce problème avec les conditions suivantes :  $H = 1$ ,  $L = 5$ ,  $\eta = 1$  et  $Q = 1$ . Pour que les résultats soient vraiment significatifs, on doit éviter d'utiliser un maillage parfaitement régulier. Nous avons donc choisi un maillage régulier (de triangles ou de rectangles) où nous avons légèrement déplacé un seul noeud (voir la figure 12.13). La figure 12.14 montre un résultat typique de solution avec l'élément  $Q_1 - Q_0$  (figure 12.7). On remarque que le champ de vitesses est acceptable mais la pression présente de fortes oscillations qui sont typiques d'un élément ne vérifiant pas la condition *inf-sup*. Au lieu d'obtenir une pression linéaire dans la direction de l'écoulement, on obtient des oscillations dues à la présence d'éléments parasites (en forme de damiers) dans l'espace  $\ker B_h^T$ .

Au contraire, l'élément *Mini* (figure 12.1) donne d'excellents résultats comme on peut le voir à la figure 12.15. La pression devrait varier entre 0 et 60 et on constate une légère erreur. Enfin, l'élément d'ordre 2  $P_2 - P_1$  (figure 12.3) donne exactement le bon résultat, à la fois pour les vitesses et la pression (voir la figure 12.16).

## 12.8 Problème de Stokes non linéaire

La vaste majorité des fluides n'entre pas dans la catégorie des fluides newtoniens, c'est-à-dire des fluides à viscosité constante. On peut observer en laboratoire divers types de comportement, notamment en élongation ou en cisaillement. La viscosité de plusieurs fluides varie en fonction du

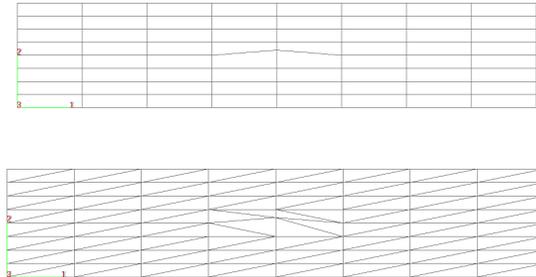


Figure 12.13 – Écoulement de Poiseuille : maillages rectangulaire et triangulaire

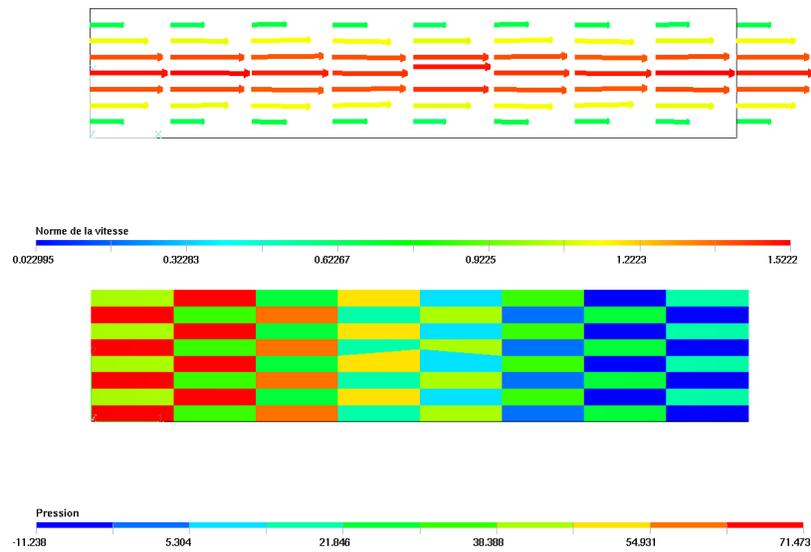


Figure 12.14 – Écoulement de Poiseuille pour l'élément  $Q_1 - Q_0$  : vitesse et pression

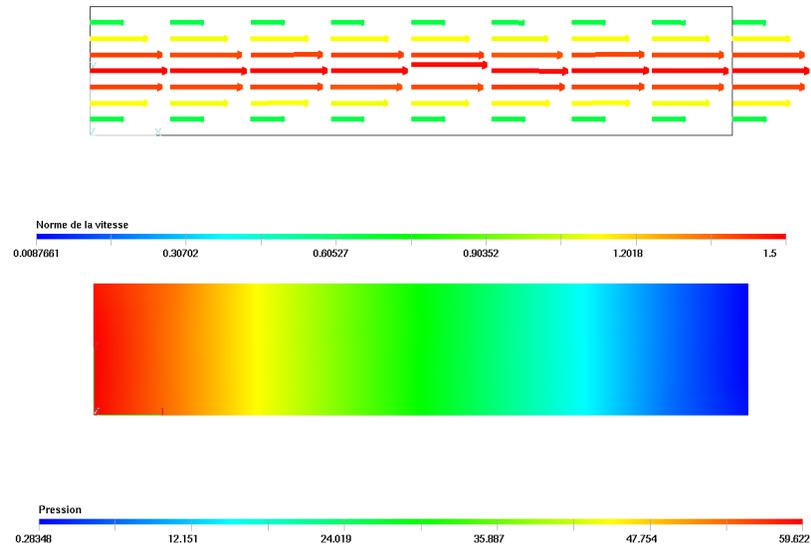


Figure 12.15 – Écoulement de Poiseuille pour l'élément Mini : vitesse et pression

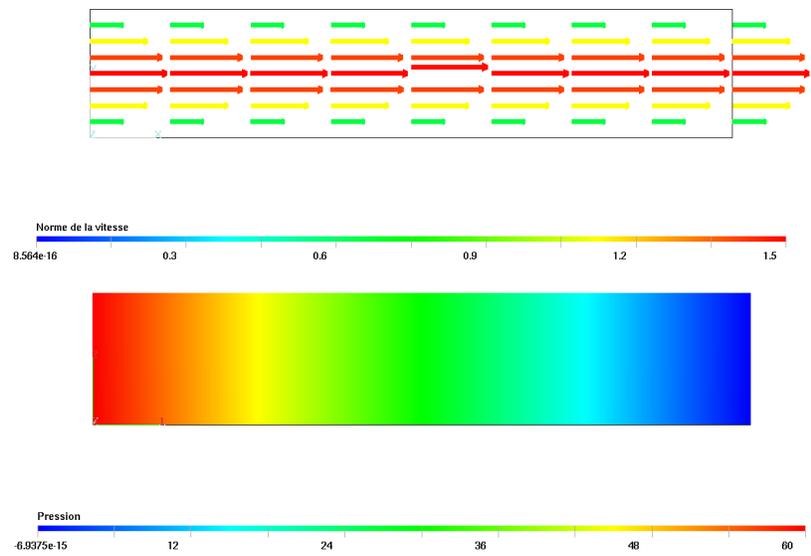


Figure 12.16 – Écoulement de Poiseuille pour l'élément  $P_2 - P_1$  : vitesse et pression

cisaillement. On se doit alors de modéliser ce comportement et d'essayer de reproduire le plus fidèlement possible les données rhéologiques de viscosité en fonction du taux de déformation. Les fluides dont la viscosité dépend du cisaillement sont dits *viscoplastiques* et les modèles correspondants ont la forme générale suivante :

$$\boldsymbol{\tau} = 2\eta(|\dot{\boldsymbol{\gamma}}(\mathbf{u})|)\dot{\boldsymbol{\gamma}}(\mathbf{u})$$

où  $|\dot{\boldsymbol{\gamma}}(\mathbf{u})|$  désigne le second invariant du tenseur  $\dot{\boldsymbol{\gamma}}(\mathbf{u})$  défini par :

$$|\dot{\boldsymbol{\gamma}}(\mathbf{u})|^2 = 2\dot{\boldsymbol{\gamma}}(\mathbf{u}) : \dot{\boldsymbol{\gamma}}(\mathbf{u}) = 2(\dot{\boldsymbol{\gamma}}(\mathbf{u}))_{ij}(\dot{\boldsymbol{\gamma}}(\mathbf{u}))_{ij}$$

La viscosité  $\eta$  est donc une fonction de  $|\dot{\boldsymbol{\gamma}}(\mathbf{u})|$  par l'intermédiaire de lois de comportement telle la loi puissance :

$$\eta = \eta_0 |\dot{\boldsymbol{\gamma}}(\mathbf{u})|^{m-1} \quad (12.20)$$

où  $m$  est l'indice de pseudoplasticité et  $\eta_0$  est la consistance. Cette loi est principalement utilisée pour décrire la viscoplasticité des métaux à chaud. On remarque cependant que la viscosité du polymère tend vers l'infini en cisaillement faible, ce qui est peu représentatif du comportement réel.

Une autre loi de comportement importante et plus réaliste est la loi de Carreau-Yasuda qui nous écrirons sous la forme :

$$\eta = \eta_0 (c_0 + \lambda^n |\dot{\boldsymbol{\gamma}}(\mathbf{u})|^n)^{\frac{m-1}{n}} \quad (12.21)$$

Dans le modèle original la constante  $c_0$  est tout simplement 1. Cette écriture nous permet cependant de voir la loi de puissance comme un cas particulier obtenu en posant  $c_0 = 0$ ,  $\lambda = 1$  et  $n = 1$ .

Le modèle de Carreau-Yasuda permet de prendre en compte la présence d'un plateau newtonien à faible taux de cisaillement. Dans ce dernier cas,  $\eta_0$  désigne toujours la consistance,  $\lambda$  est un temps de relaxation et  $m$  et  $n$  des indices de pseudoplasticité. En particulier, si  $n = 2$  on retrouve le modèle classique de Carreau.

Nous présentons maintenant la formulation variationnelle dans le cas des fluides viscoplastiques. Nous supposons donc que la viscosité du polymère est donnée par la loi de Carreau-Yasuda sous la forme 12.21. La formulation variationnelle élémentaire du problème de Stokes devient dans ce cas :

$$\left\{ \begin{array}{l} \int_K 2\eta_0(c_0 + \lambda^n |\dot{\boldsymbol{\gamma}}(\mathbf{u})|^n)^{\frac{m-1}{n}} \dot{\boldsymbol{\gamma}}(\mathbf{u}) : \dot{\boldsymbol{\gamma}}(\mathbf{w}) \, dv - \int_K p \boldsymbol{\nabla} \cdot \mathbf{w} \, dv = \int_K \mathbf{r}^K \cdot \mathbf{w} \, dv + \int_{\partial K} \mathbf{t}^K \cdot \mathbf{w} \, ds \\ \int_K q \boldsymbol{\nabla} \cdot \mathbf{u} \, dv = 0 \end{array} \right.$$

De toute évidence, le problème est maintenant non linéaire et on peut appliquer les techniques vues au chapitre 8 et notamment la méthode de Newton. Bien entendu, on suppose connues la fonction  $\mathbf{r}^K$  de même que la traction  $\mathbf{t}^K$ .

Partant d'une solution approximative  $(\mathbf{u}_0, p_0)$ , on cherche une correction  $(\boldsymbol{\delta}_u, \delta_p)$  de sorte que  $(\mathbf{u}_0 + \boldsymbol{\delta}_u, p_0 + \delta_p)$  soit une solution du problème non linéaire ou du moins une meilleure approximation. Remplaçant  $\mathbf{u}$  et  $p$  par  $\mathbf{u}_0 + \boldsymbol{\delta}_u$  et  $p_0 + \delta_p$  on trouve :

$$\left\{ \begin{array}{l} \int_K 2\eta_0(c_0 + \lambda^n |\dot{\gamma}(\mathbf{u}_0 + \boldsymbol{\delta}_u)|^n)^{\frac{m-1}{n}} \dot{\gamma}(\mathbf{u}_0 + \boldsymbol{\delta}_u) : \dot{\gamma}(\mathbf{w}) \, dv - \int_K (p_0 + \delta_p) \nabla \cdot \mathbf{w} \, dv = \\ \int_K \mathbf{r}^K \cdot \mathbf{w} \, dv + \int_{\partial K} \mathbf{t}^K \cdot \mathbf{w} \, ds \\ \int_K q \nabla \cdot (\mathbf{u}_0 + \boldsymbol{\delta}_u) \, dv = 0 \end{array} \right.$$

ou encore :

$$\left\{ \begin{array}{l} \int_K 2\eta_0(c_0 + \lambda^n |\dot{\gamma}(\mathbf{u}_0 + \boldsymbol{\delta}_u)|^n)^{\frac{m-1}{n}} \dot{\gamma}(\mathbf{u}_0 + \boldsymbol{\delta}_u) : \dot{\gamma}(\mathbf{w}) \, dv - \int_K \delta_p \nabla \cdot \mathbf{w} \, dv = \\ \int_K p_0 \nabla \cdot \mathbf{w} \, dv + \int_K \mathbf{r}^K \cdot \mathbf{w} \, dv + \int_{\partial K} \mathbf{t}^K \cdot \mathbf{w} \, ds \\ \int_K q \nabla \cdot \boldsymbol{\delta}_u \, dv = - \int_K q \nabla \cdot \mathbf{u}_0 \, dv \end{array} \right.$$

Il reste alors un dernier terme à linéariser. Pour ce faire on utilise la relation 8.9 et la dérivée de Gâteaux 8.10. On obtient ainsi le problème linéarisé :

$$\left\{ \begin{array}{l} a_{\mathbf{u}_0}(\boldsymbol{\delta}_u, \mathbf{w}) - \int_K \delta_p \nabla \cdot \mathbf{w} \, dv = - \int_K \mathbf{R}((\mathbf{u}_0, p_0), \mathbf{w}) \, dv \\ \int_K q \nabla \cdot \boldsymbol{\delta}_u \, dv = - \int_K q \nabla \cdot \mathbf{u}_0 \, dv \end{array} \right.$$

où la forme bilinéaire  $a_{\mathbf{u}_0}$  et le résidu  $R((\mathbf{u}_0, p_0), \mathbf{w})$  sont définis par :

$$\begin{aligned} a_{\mathbf{u}_0}(\boldsymbol{\delta}_u, \mathbf{w}) &= \int_K \left( 2\eta_0(c_0 + \lambda^n |\dot{\gamma}(\mathbf{u}_0)|^n)^{\frac{m-1}{n}} \dot{\gamma}(\boldsymbol{\delta}_u) : \dot{\gamma}(\mathbf{w}) \right) \, dv + \\ &\quad \int_K K_0 \left( |\dot{\gamma}(\mathbf{u}_0)|^{n-2} (c_0 + \lambda^n |\dot{\gamma}(\mathbf{u}_0)|^n)^{\frac{m-n-1}{n}} (\dot{\gamma}(\mathbf{u}_0) : \dot{\gamma}(\boldsymbol{\delta}_u)) (\dot{\gamma}(\mathbf{u}_0) : \dot{\gamma}(\mathbf{w})) \right) \, dv \\ R((\mathbf{u}_0, p_0), \mathbf{w}) &= \int_K 2\eta_0(c_0 + \lambda^n |\dot{\gamma}(\mathbf{u}_0)|^n)^{\frac{m-1}{n}} \dot{\gamma}(\mathbf{u}_0) : \dot{\gamma}(\mathbf{w}) \, dv - \int_K p_0 \nabla \cdot \mathbf{w} \, dv - \int_K \mathbf{r}^K \cdot \mathbf{w} \, dv \\ &\quad - \int_{\partial K} \mathbf{t}^K \cdot \mathbf{w} \, ds \end{aligned} \tag{12.22}$$

où on a posé  $K_0 = 4\eta_0(m-1)\lambda^n$

## 12.8.1 Résultats numériques : cas viscoplastique

### Écoulement de Poiseuille

On considère encore un écoulement de Poiseuille mais cette fois avec une loi puissance 12.20 pour laquelle la viscosité diminue avec le cisaillement. Ce cisaillement se produisant principalement



Figure 12.17 – Écoulement de Poiseuille : loi puissance avec  $m = 1$ ,  $m = 0,5$  et  $m = 0,3$

aux parois solides, la viscosité  $\gamma$  est plus faible qu’au centre du canal. On remarque à la figure 12.17 la différence de comportement entre les cas newtoniens ( $m = 1$ ) et les cas non newtoniens  $m = 0,7$  et  $m = 0,4$ . Le profil de vitesse en sortie s’aplatit lorsque  $m$  diminue et on le voit encore mieux à la figure 12.18 qui montre une coupe du profil de vitesse  $u_1$  en sortie sur l’axe  $x_1 = L$ . On y voit clairement l’aplatissement du profil et on remarque également que la surface sous la courbe (le débit) reste la même.

Le comportement de la pression est bien caractéristique d’un fluide viscoplastique (voir la figure 12.19). Pour un débit donné, la perte de charge c.-à-d. la différence de pression entre l’entrée et la sortie du canal diminue fortement avec  $m$ . En effet, puisque la viscosité diminue avec  $m$ , il faut pousser moins fort pour assurer un même débit. Cette propriété est fort appréciée dans bon nombre de procédés de mise en forme des polymères comme l’extrusion (voir Agassant [1]).

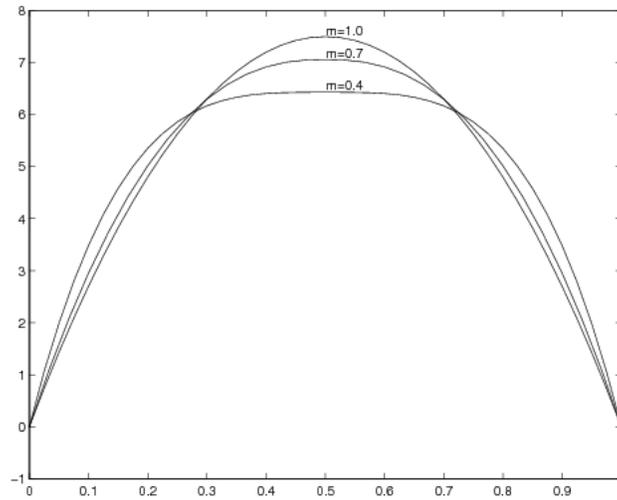


Figure 12.18 – Profil de vitesse en sortie du canal

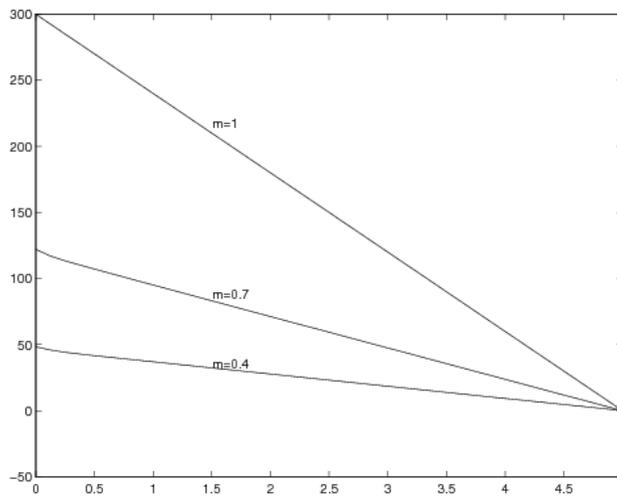


Figure 12.19 – Perte de charge dans le canal

### Écoulement dans une filière d'extrusion

On considère dans ce problème une filière d'extrusion utilisée pour la production de poutre en I constituée d'un mélange de polymère et de fibre de bois. La section initiale est de forme circulaire et prend graduellement la forme en I désirée. Le problème a été adimensionnalisé de sorte que le rayon de la section circulaire de la filière soit 1. La longueur de la filière est d'environ 5,7 suivant l'axe des  $z$ . Le mélange bois-plastique est réputé viscoplastique avec les valeurs suivantes des paramètres du modèle de Carreau :  $\eta_0 = 1$ ,  $\lambda = 0,1588$ ,  $n = 2$  et  $m = 0,3$ .

La valeur de  $m$  étant assez loin du cas newtonien ( $m = 1$ ), la résolution du problème a nécessité plusieurs étapes. Nous avons d'abord résolu en prenant  $m = 1$  pour obtenir une solution de départ du problème non linéaire. par la suite, nous avons résolu avec  $m = 0,8$ ,  $m = 0,5$  et  $m = 0,3$ , toujours en prenant la solution précédente comme point de départ de la nouvelle résolution. On a utilisé un élément  $P_2 - P_1$  c.-à-d. quadratique en vitesse et linéaire en pression (voir la figure 12.4). Les conditions aux limites sont les suivantes :

$$\left\{ \begin{array}{ll} \mathbf{u} &= \left( 0, 0, \frac{(3m+1)}{(2m+2)}(1 - r^{(1+m)/m}) \right) \quad \text{où } r = (x^2 + y^2)^{1/2} \quad \text{sur la section circulaire d'entrée} \\ \mathbf{u} &= (0, 0, 0) \quad \text{sur les parois} \\ \boldsymbol{\sigma} \cdot \mathbf{n} &= (0, 0, 0) \quad \text{sur la section de sortie} \end{array} \right.$$

Le profil de vitesse en entrée correspond à un écoulement de Poiseuille dans un tube de section circulaire pour un fluide en loi de puissance. Ce profil est légèrement inexact pour le modèle de Carreau.

Le maillage est illustré à la figure 12.20 de même que la pression et le module de vitesse à la sortie de la filière. En ce qui concerne la pression, on remarque qu'elle est à peu près constante dans la partie circulaire de la filière. On observe ensuite une chute rapide jusqu'en sortie où la pression s'annule en raison de la condition aux limites. Le module de vitesse est évidemment nul sur les parois de la filière. En sortie, on peut facilement voir l'accélération provoquée par la forme en I de la filière. En effet, l'écoulement est favorisé dans les parties plus larges.

## 12.9 Les équations de Navier-Stokes

Les fluides viscoplastiques sont généralement très visqueux, leur viscosité étant souvent de l'ordre de  $10^5 \text{ Pa} \cdot \text{s}$ . À l'autre extrémité du spectre, on trouve les fluides peu visqueux comme l'eau et l'air. Il va de soi que ces fluides particuliers sont aussi d'une grande importance en hydraulique et en aéronautique, bien que dans ce dernier cas, il faut tenir compte de la compressibilité.

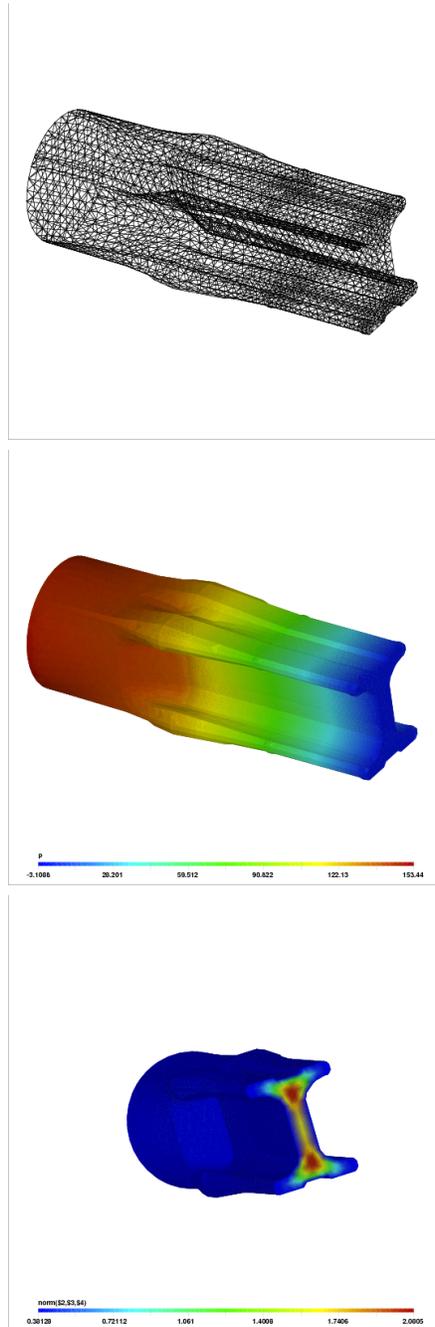


Figure 12.20 – Maillage, perte de charge et module de vitesse en sortie de filière

Dans ce dernier cas, les équations d'équilibre sont modifiées de la manière suivante :

$$\begin{cases} -\nabla \cdot (2\eta\dot{\gamma}(\mathbf{u})) + \nabla p &= \rho \mathbf{r} - \rho \left( \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right) \\ \nabla \cdot \mathbf{u} &= 0 \end{cases} \quad (12.23)$$

Le deuxième terme de droite en est un d'inertie, qui est négligé dans le cas où la viscosité est grande (et où par le fait même le terme visqueux est dominant). Notons de plus la forme spécifique du terme source  $\rho \mathbf{r}$  où  $\mathbf{r}$  désigne une force par unité de masse (dont les unités sont des N/kg). Nous négligerons dans un premier temps la dérivée en temps pour ne considérer que le cas stationnaire. Même si la viscosité est constante, ce système est non linéaire.

On retrouve souvent les équations de Navier-Stokes sous forme adimensionnelle. Pour ce faire, on choisit une longueur  $L_0$ , une vitesse  $U_0$ , une pression  $p_0$  et une force massique  $r_0$ , dites de référence et qui dépendent de la géométrie et des conditions du problème. On introduit ensuite les variables adimensionnelles  $\tilde{\mathbf{x}} = \mathbf{x}/L_0$ ,  $\tilde{\mathbf{u}} = \mathbf{u}/U_0$ ,  $\tilde{p} = p/P_0$  et  $\tilde{\mathbf{r}} = \mathbf{r}/R_0$  et en faisant ce changement de variables, on obtient les équations adimensionnelles :

$$\begin{cases} -\frac{1}{Re} \tilde{\nabla} \cdot (2\tilde{\gamma}(\tilde{\mathbf{u}})) + \tilde{\nabla} \tilde{p} + (\tilde{\mathbf{u}} \cdot \tilde{\nabla}) \tilde{\mathbf{u}} &= \tilde{\mathbf{r}} \\ \tilde{\nabla} \cdot \tilde{\mathbf{u}} &= 0 \end{cases} \quad (12.24)$$

où on a introduit le nombre adimensionnel dit *de Reynolds* qui est défini par :

$$Re = \frac{U_0 L_0 \rho}{\eta}$$

et où on a choisi  $P_0 = \rho U_0^2$  et  $R_0 = P_0/(\rho L_0)$  comme pression et force massique de référence. Les variables identifiées par le symbole  $\sim$  sont adimensionnelles (n'ont pas d'unités). Dans ce qui suit, nous laisseront tomber ce symbole pour éviter d'alourdir la notation.

On peut, ici encore, appliquer la méthode de Newton, bien qu'il existe d'autres possibilités. Il n'y a dans ce cas qu'un terme à linéariser. Par les techniques que nous avons vues, on obtient ainsi le problème linéarisé :

$$\begin{cases} a_{\mathbf{u}_0}(\delta_{\mathbf{u}}, \mathbf{w}) - \int_K \delta_p \nabla \cdot \mathbf{w} \, dv &= - \int_K \mathbf{R}((\mathbf{u}_0, p_0), \mathbf{w}) \, dv \\ \int_K q \nabla \cdot \delta_{\mathbf{u}} \, dv &= - \int_K q \nabla \cdot \mathbf{u}_0 \, dv \end{cases}$$

où la forme bilinéaire  $a_{\mathbf{u}_0}$  et le résidu  $R((\mathbf{u}_0, p_0), \mathbf{w})$  sont définis par :

$$\begin{aligned}
 a_{\mathbf{u}_0}(\boldsymbol{\delta}_u, \mathbf{w}) &= \int_K \left( \frac{2}{Re} \dot{\gamma}(\boldsymbol{\delta}_u) : \dot{\gamma}(\mathbf{w}) + ((\mathbf{u}_0 \cdot \nabla) \boldsymbol{\delta}_u) \cdot \mathbf{w} + ((\boldsymbol{\delta}_u \cdot \nabla) \mathbf{u}_0) \cdot \mathbf{w} \right) dv \\
 R((\mathbf{u}_0, p_0), \mathbf{w}) &= \int_K \frac{2}{Re} \dot{\gamma}(\mathbf{u}_0) : \dot{\gamma}(\mathbf{w}) + (\mathbf{u}_0 \cdot \nabla \mathbf{u}_0) \cdot \mathbf{w} dv - \int_K p_0 \nabla \cdot \mathbf{w} dv - \int_K \mathbf{r}^K \cdot \mathbf{w} dv \\
 &\quad - \int_{\partial K} \mathbf{t}^K \cdot \mathbf{w} ds
 \end{aligned}
 \tag{12.25}$$

Différentes variantes de la méthode des points fixes sont également envisageables.

### 12.9.1 Résultats numériques : cas Navier-Stokes

#### Écoulement autour d'un cylindre : cas stationnaire

On considère l'écoulement d'un fluide autour d'un cylindre circulaire de diamètre 1 et centré en  $(0, 0)$ . Pour éviter une influence indue des conditions aux limites, on fera le calcul dans une géométrie assez vaste c.-à-d.  $-10 \leq x \leq 25$  et  $-10 \leq y \leq 10$ . La raison de ce choix est que les conditions aux limites sont mal connues en aval du cylindre. Par contre, sur la section d'entrée ( $x = -10$ ) on imposera un profil de vitesse uniforme  $\mathbf{u} = (1, 0)$  de même que sur les frontières (fictives) supérieures et inférieures. On suppose que ces deux frontières sont suffisamment éloignées pour que le profil uniforme ne soit pas perturbé. Enfin, on imposera une condition naturelle nulle en sortie  $\boldsymbol{\sigma} \cdot \mathbf{n} = (0, 0)$ . Cette dernière condition est encore ici douteuse, spécialement dans le cas instationnaire. Les conditions aux limites sont présentées à la figure 12.21. Le maillage de 14700 éléments de Taylor-Hood ( $P_2 - P_1$ ) est présenté à la figure 12.22 de même que les lignes de courant à  $Re = 1$ . On constate que loin du cylindre, les lignes de courant sont peu perturbées, ce qui justifie *a posteriori* le choix des conditions aux limites.

Pour bien cerner l'influence du nombre de Reynolds sur la solution, on retrace les lignes de courant immédiatement au voisinage du cylindre. C'est ce qui est illustré à la figure 12.23 où on peut constater les différences entre les solutions à  $Re = 1, 50$  et  $100$ . À très faibles nombres de Reynolds, les lignes de courant contournent le cylindre avec un minimum de perturbations. Par contre, en augmentant sa valeur, la couche limite se détache et on observe la formation de zones de recirculation (tourbillons) dans le sillage.

#### Écoulement autour d'un cylindre : cas instationnaire

Lorsque le nombre de Reynolds atteint environ  $Re = 45$ , l'écoulement devient instationnaire et il se développe à l'arrière du cylindre une instabilité connue sous le nom d'allée de von Karman. Ce phénomène est très fréquemment rencontré dans la nature et c'est ce qui fait par exemple louvoyer un drapeau dans le vent, le mât au bout duquel flotte le drapeau faisant office de cylindre. Le passage d'un écoulement stationnaire à un écoulement instationnaire correspond à ce que les mathématiciens appellent une bifurcation de Hopf.

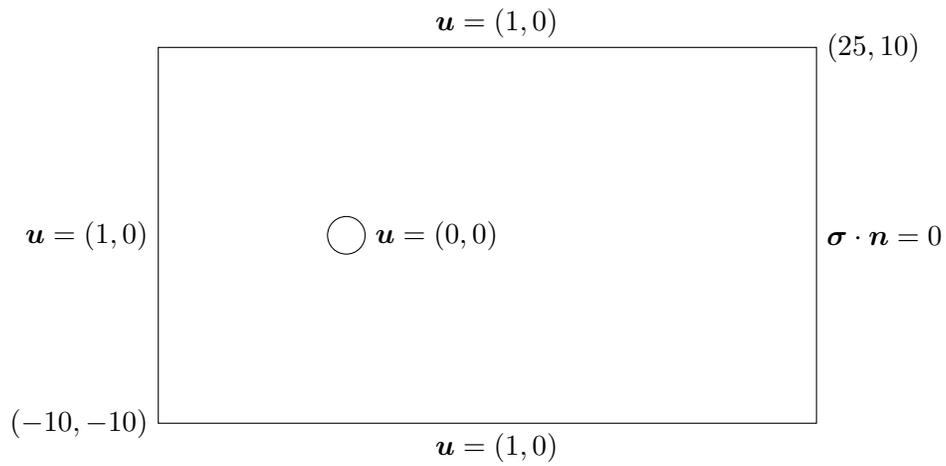


Figure 12.21 – Écoulement autour d'un cylindre : géométrie et conditions aux limites

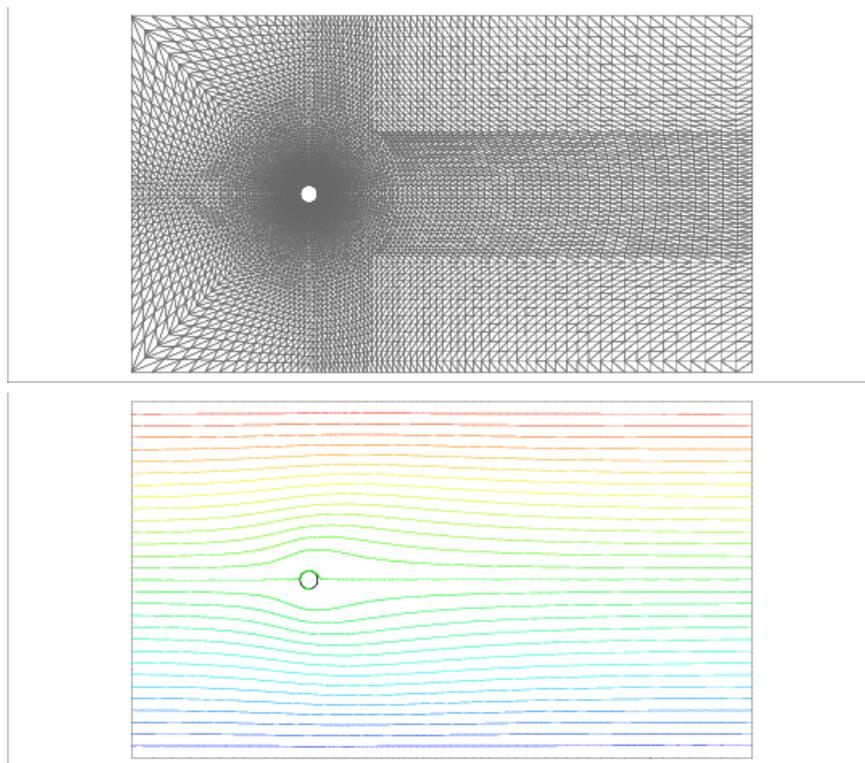


Figure 12.22 – Maillage et portrait global des lignes de courant à  $Re = 1$

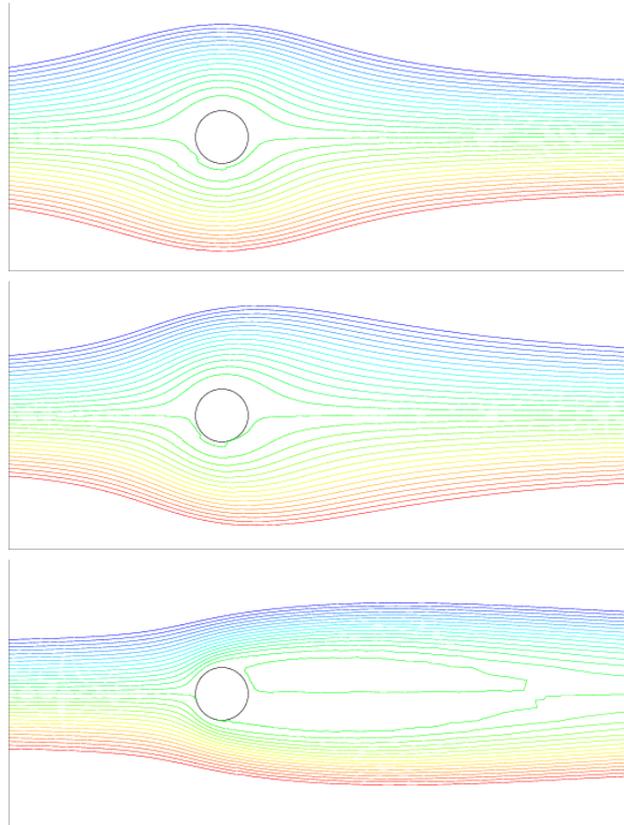


Figure 12.23 – Lignes de courant derrière le cylindre à  $Re = 1, 50$  et  $100$

Nous présentons maintenant des résultats obtenus pour un nombre de Reynolds de 100, nombre bien au-delà du nombre critique. Nous avons utilisé une différence arrière implicite d'ordre 2 pour la dérivée en temps ainsi qu'un pas de temps de 0,1.

La figure 12.24 donne les détails de l'écoulement derrière le cylindre sur près d'une période. En effet, on peut vérifier que l'écoulement est bien périodique de période 5,9, en accord notamment avec Engleman et Jamnia [16]. Puisque  $\Delta t = 0,1$ , une soixantaine de pas de temps couvre donc une période d'oscillations. Il est à noter qu'au départ, un grand nombre de pas de temps sont nécessaires pour éliminer les effets transitoires et pour que l'écoulement périodique s'établisse de manière définitive.

On trouvera un certain nombre d'exemples de modélisation d'écoulements instationnaires dans Fortin et al. [18]. On y observera des bifurcations de Hopf ainsi que ce qu'il est convenu d'appeler des cascades de dédoublements de périodes, correspondant à de nouvelles bifurcations.

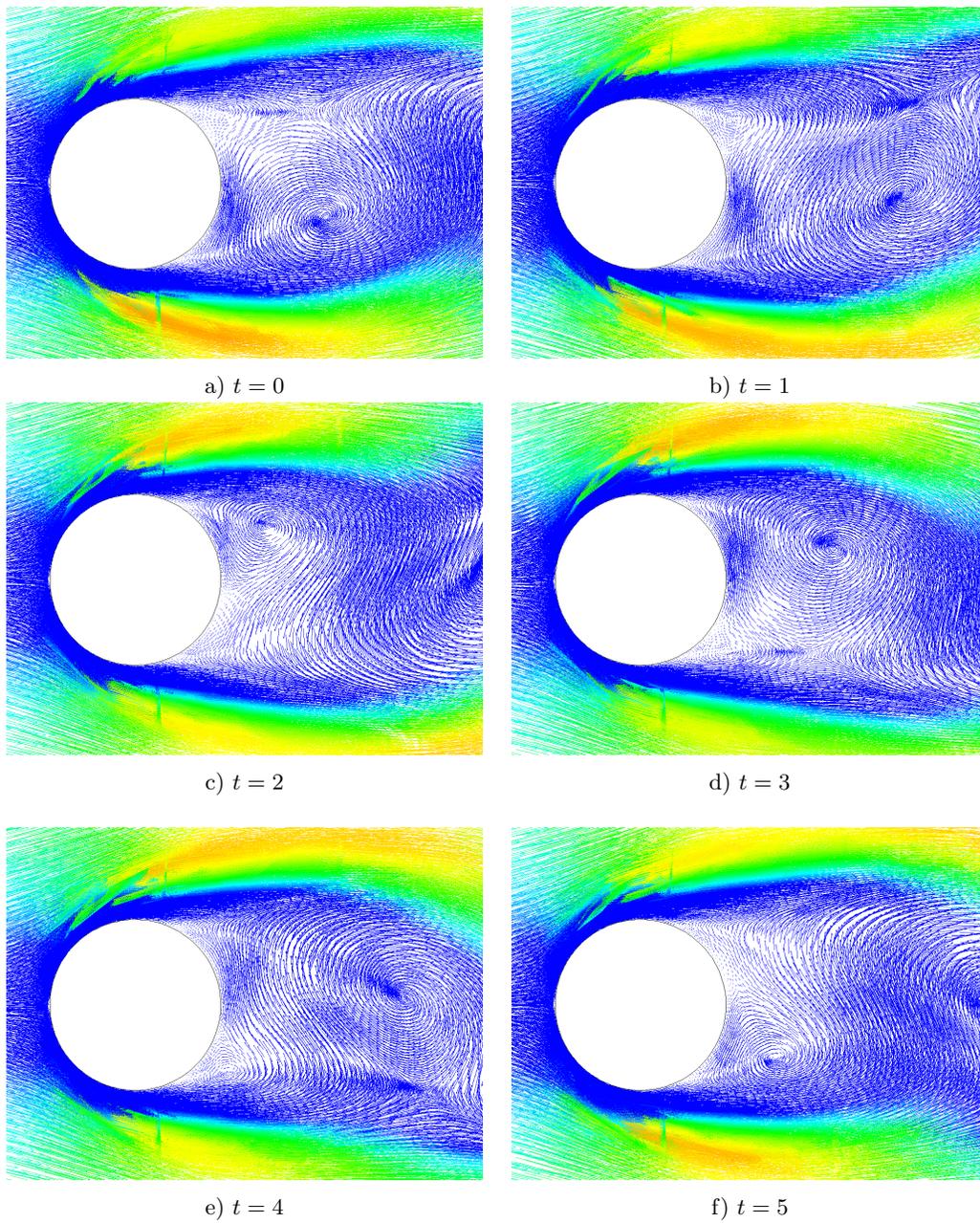


Figure 12.24 – Champ de vitesse derrière le cylindre à  $Re = 100$  sur une période

## 12.10 Exercices

1. Montrer que l'opérateur  $B$  défini par la relation :

$$\langle B\mathbf{w}, q \rangle = b(\mathbf{w}, q)$$

est bien linéaire et continu sur  $V$ .

2. On va démontrer que pour le problème de Stokes, une discrétisation linéaire par éléments en vitesse et en pression ( $P_1 - P_1$ ) ne vérifie pas la condition de Brezzi ou plus précisément que le noyau de l'opérateur  $B_h^T$  contient des pressions parasites. On considère encore le maillage de la figure 6.15 ainsi que 3 nombres  $a, b$  et  $c$  tels que  $a + b + c = 0$  c.-à-d. de moyenne nulle.

- a) Reproduire le maillage de la figure 6.15 et assigner à chaque noeud du maillage l'une des valeurs  $a, b$  ou  $c$  de sorte que la moyenne des valeurs nodales dans tous les éléments soit nulle.
- b) On considère maintenant la fonction linéaire par élément  $p_h(\mathbf{x})$  prenant les valeurs nodales obtenues en (a). Montrer à l'aide d'une quadrature de Hammer à un point que si les valeurs nodales sont de moyenne nulle dans un élément, alors :

$$\int_K p_h(\mathbf{x}) dv = 0 \quad \forall K$$

- c) Montrer alors que si les composantes du vecteur vitesse sont linéaires par élément, alors la fonction  $p_h$  construite précédemment est dans le noyau de l'opérateur  $B_h^T$  c.-à-d.

$$b(\mathbf{w}_h, p_h) = \int_{\Omega} p_h \nabla \cdot \mathbf{w}_h dv = 0 \quad \forall \mathbf{w}_h \in V_h$$

3. En partant des équations de Navier-Stokes dimensionnelles 12.23, obtenir la forme adimensionnelle 12.24 en choisissant en particulier  $P_0 = \rho U_0^2$  et  $R_0 = P_0 / (\rho L_0)$  comme pression et force massique de référence.

## Chapitre 13

# Formulations mixtes en élasticité linéaire

Au chapitre 11, nous avons rencontré ce qu'il est convenu d'appeler une *formulation déplacement* du problème d'élasticité linéaire isotrope en ce sens que la seule inconnue du problème est le vecteur déplacement  $\mathbf{u}$ . Nous avons vu au chapitre 12 une formulation mixte du problème de Stokes (en vitesse et pression) et constaté l'importance de choisir des approximations consistantes en vitesse et pression. Nous revenons dans ce chapitre sur les problèmes d'élasticité linéaire, mais en formulation mixte (en déplacement et pression). De telles formulations mixtes deviennent nécessaires lorsque le coefficient de Poisson  $\nu$  approche 0,5 et que le matériau devient incompressible. Le système obtenu est alors très proche du système de Stokes. Nous mettrons également en évidence la faiblesse de la formulation déplacement dans ce cas.

### 13.1 Cas isotrope

Rappelons les équations d'équilibre :

$$-\nabla \cdot \boldsymbol{\sigma} = \mathbf{r}$$

et l'expression du tenseur des contraintes dans le cas isotrope :

$$\boldsymbol{\sigma} = \lambda(\nabla \cdot \mathbf{u})\mathbf{I} + 2\mu\boldsymbol{\gamma}(\mathbf{u}) \quad (13.1)$$

où  $\mathbf{I}$  est le tenseur identité et  $\boldsymbol{\gamma}(\mathbf{u})$  est le tenseur de déformation défini à la relation 11.2. Notons que ceci est un cas particulier de la relation :

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\gamma}(\mathbf{u}) \stackrel{\text{déf.}}{=} \mathcal{C}_{ijkl}(\boldsymbol{\gamma}(\mathbf{u}))_{kl} = \mathcal{C}_{ijkl}\gamma_{kl}$$

où on a posé  $\gamma_{ij} = (\boldsymbol{\gamma}(\mathbf{u}))_{ij}$  et où on a utilisé la convention de somme sur les indices répétés.  $\mathbf{C}$  est le tenseur d'élasticité (du quatrième ordre). Dans le cas isotrope, le tenseur d'élasticité est défini

par la relation 11.9 que nous rappelons :

$$\mathcal{C}_{ijkl} = \lambda I_{ij} I_{kl} + \mu (I_{ik} I_{jl} + I_{il} I_{jk}) \quad (13.2)$$

Rappelons que  $\lambda$  et  $\mu$  sont les coefficients de Lamé (voir la relation 11.10). On introduit également le coefficient de compressibilité  $k$  (« bulk modulus » en anglais) défini par :

$$k = \frac{E}{3(1 - 2\nu)}$$

et on vérifie facilement que :

$$k = \frac{(3\lambda + 2\mu)}{3} \quad \text{ou encore} \quad \lambda - k = -\frac{2\mu}{3} \quad (13.3)$$

On définit la pression par :

$$p = -\frac{\text{tr}(\boldsymbol{\sigma})}{3}$$

Puisque la trace de  $\boldsymbol{\sigma}$  est :

$$\text{tr}(\boldsymbol{\sigma}) = \text{tr}(\lambda(\nabla \cdot \mathbf{u})\mathbf{I} + 2\mu\boldsymbol{\gamma}(\mathbf{u})) = \lambda(\nabla \cdot \mathbf{u})\text{tr}(\mathbf{I}) + 2\mu \text{tr}(\boldsymbol{\gamma}(\mathbf{u})) = (3\lambda + 2\mu)(\nabla \cdot \mathbf{u}) = 3k(\nabla \cdot \mathbf{u})$$

on en conclut que :

$$p = -k(\nabla \cdot \mathbf{u})$$

Il est alors courant (voir les références Bathe [4] ainsi que Zienkiewicz et Taylor [39] par exemple) de décomposer le tenseur des contraintes  $\boldsymbol{\sigma}$  en une partie de trace nulle que l'on appelle le *déviateur* de  $\boldsymbol{\sigma}$  (notée  $\boldsymbol{\tau}$ ) et une partie dite *sphérique* faisant intervenir la pression. Selon Bathe [4], on pose alors :

$$\boldsymbol{\tau} = \boldsymbol{\sigma} - \frac{\text{tr}(\boldsymbol{\sigma})}{3}\mathbf{I} = \boldsymbol{\sigma} - k(\nabla \cdot \mathbf{u})\mathbf{I} = 2\mu\boldsymbol{\gamma}(\mathbf{u}) - \frac{2\mu}{3}(\nabla \cdot \mathbf{u})\mathbf{I}$$

Par construction, on a  $\text{tr}(\boldsymbol{\tau}) = 0$  et de plus :

$$\boldsymbol{\sigma} = \boldsymbol{\tau} + \frac{\text{tr}(\boldsymbol{\sigma})}{3}\mathbf{I} = \boldsymbol{\tau} - p\mathbf{I}$$

ou encore :

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\gamma}(\mathbf{u}) - \frac{2\mu}{3}(\nabla \cdot \mathbf{u})\mathbf{I} - p\mathbf{I} \quad (13.4)$$

où on a utilisé les relations 13.1 et 13.3. L'équation d'équilibre devient alors un système de 2 équations faisant intervenir le déplacement  $\mathbf{u}$  et la pression  $p$  :

$$\begin{cases} -\nabla \cdot \left[ 2\mu\boldsymbol{\gamma}(\mathbf{u}) - \frac{2\mu}{3}(\nabla \cdot \mathbf{u})\mathbf{I} - p\mathbf{I} \right] = \mathbf{r} \\ -\frac{1}{k}p - (\nabla \cdot \mathbf{u}) = 0 \end{cases}$$

Nous devons donc obtenir des discrétisations par éléments finis de cette formulation mixte. Multipliant la première équation par une fonction test  $\mathbf{w}$  et la seconde par  $q$  et après intégration par parties, on trouve :

$$\begin{aligned} \int_{\Omega} \left( 2\mu\gamma(\mathbf{u}) : \gamma(\mathbf{w}) - \frac{2\mu}{3}(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{w}) \right) dv - \int_{\Omega} p\nabla \cdot \mathbf{w} dv &= \int_{\Omega} \mathbf{r} \cdot \mathbf{w} dv + \int_{\Gamma} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \mathbf{w} d\Gamma \\ - \int_{\Omega} q\nabla \cdot \mathbf{u} dv - \frac{1}{k} \int_{\Omega} pq dv &= 0 \end{aligned} \quad (13.5)$$

La condition essentielle est, comme pour le problème de Stokes, l'imposition du déplacement  $\mathbf{u}$  (ou de l'une de ses composantes  $u_i$ ) et la condition naturelle consiste à imposer  $\boldsymbol{\sigma} \cdot \mathbf{n}$  (ou de l'une de ses composantes  $(\boldsymbol{\sigma} \cdot \mathbf{n})_i$ ).

L'analyse numérique du système 13.5 ressemble beaucoup à celle du problème de Stokes. En fait, lorsque le coefficient de Poisson s'approche de 0,5, le coefficient de compressibilité  $k$  tend vers l'infini, ce qui annule le deuxième terme de la deuxième équation du système. On retombe alors sur le problème de Stokes à la différence près du terme :

$$\frac{2\mu}{3} \int_{\Omega} (\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{w}) dv$$

C'est donc dans le cas où  $\nu \simeq 0,5$  que les méthodes mixtes pour l'élasticité linéaire prennent toute leur importance et se distinguent avantageusement des formulations en déplacement seulement.

Nous ne referons pas toute l'analyse de cette méthode mixte puisqu'elle est tout-à-fait similaire à celle du problème de Stokes. Il y a toutefois de petites différences qu'il faut souligner. Le problème d'élasticité linéaire mixte pour un matériau isotrope s'écrit :

$$\begin{cases} a(\mathbf{u}, \mathbf{w}) + b(\mathbf{w}, p) = (\mathbf{r}, \mathbf{w}) & \forall \mathbf{w} \in V \\ b(\mathbf{u}, q) - c(p, q) = 0 & \forall q \in Q \end{cases} \quad (13.6)$$

où :

$$\begin{aligned} a(\mathbf{u}, \mathbf{w}) &= \int_{\Omega} \left( 2\mu\gamma(\mathbf{u}) : \gamma(\mathbf{w}) - \frac{2\mu}{3}(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{w}) \right) dv \\ b(\mathbf{w}, q) &= - \int_{\Omega} q\nabla \cdot \mathbf{w} dv \\ c(p, q) &= \frac{1}{k} \int_{\Omega} pq dv \end{aligned}$$

◆ **Remarque 13.1**

On montre facilement que :

$$\begin{aligned}
 a(\mathbf{u}, \mathbf{w}) &= \int_{\Omega} \left( 2\mu\boldsymbol{\gamma}(\mathbf{u}) : \boldsymbol{\gamma}(\mathbf{w}) - \frac{2\mu}{3}(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{w}) \right) dv \\
 &= 2\mu \int_{\Omega} \left( \boldsymbol{\gamma}(\mathbf{u}) - \frac{\nabla \cdot \mathbf{u}}{3}\mathbf{I} \right) : \boldsymbol{\gamma}(\mathbf{w}) dv \\
 &= 2\mu \int_{\Omega} \left( \boldsymbol{\gamma}(\mathbf{u}) - \frac{\nabla \cdot \mathbf{u}}{3}\mathbf{I} \right) : \left( \boldsymbol{\gamma}(\mathbf{w}) - \frac{\nabla \cdot \mathbf{w}}{3}\mathbf{I} \right) dv
 \end{aligned}$$

ce qui entraîne que  $a(\mathbf{w}, \mathbf{w})$  est une quantité qui est toujours positive.  $\blacklozenge$

### ◆ Remarque 13.2

On peut obtenir une formulation mixte des problèmes d'élasticité à partir de la relation 13.4 en posant simplement  $\tilde{p} = -\lambda\nabla \cdot \mathbf{u}$ . On obtient alors une forme plus simple que l'on rencontre fréquemment (voir Brezzi et Fortin [7] par exemple) :

$$\begin{aligned}
 -\nabla \cdot (2\mu\boldsymbol{\gamma}(\mathbf{u}) - \tilde{p}\mathbf{I}) &= \mathbf{r} \\
 -\frac{\tilde{p}}{\lambda} + \nabla \cdot \mathbf{u} &= 0
 \end{aligned}$$

Il faut cependant noter que la variable  $\tilde{p}$  n'est plus la partie sphérique du tenseur des contraintes et qu'elle est différente de  $p$ .  $\blacklozenge$

### ★ Théorème 13.1 (Existence et unicité du problème continu)

Sous les conditions suivantes :

- la forme bilinéaire  $a$  vérifie  $a(\mathbf{w}, \mathbf{w}) \geq 0 \forall \mathbf{w} \in V$  et est coercive sur  $\ker B$  :

$$a(\mathbf{w}, \mathbf{w}) \geq \alpha \|\mathbf{w}\|_{1,\Omega}^2 \quad \forall \mathbf{w} \in \ker B \quad (13.7)$$

- la forme bilinéaire  $c$  est symétrique et telle que :  $c(q, q) \geq 0 \forall q \in Q$  ;
- il existe une constante positive  $\beta$  telle que la forme bilinéaire  $b$  vérifie :

$$\inf_{q \in Q} \sup_{\mathbf{w} \in V} \frac{b(\mathbf{w}, q)}{\|q\|_{0,\Omega} \|\mathbf{w}\|_{1,\Omega}} \geq \beta \quad (13.8)$$

alors la solution  $(\mathbf{u}, p)$  du problème d'élasticité linéaire isotrope 13.6 est unique dans  $V \times Q$ .

**Démonstration** : Voir Brezzi-Fortin [7].

Contrairement au problème de Stokes, la forme bilinéaire  $a$  n'est coercive que sur  $\ker B$ . La condition *inf-sup* est par contre la même. ★

◆ **Remarque 13.3**

La présence de la forme bilinéaire  $c$  permet d'assurer l'unicité de la pression  $p$ . ◆

## 13.2 Cas anisotrope

Le cas général des matériaux anisotropes est plus difficile mais est tout de même d'une grande importance. De nombreux matériaux ne sont pas isotropes. Le bois, par exemple est un matériau orthotrope et ses propriétés ne sont de toute évidence pas les mêmes dans toutes les directions.

L'analyse qui suit couvre tous les cas, y compris bien sûr les matériaux isotropes. Nous suivrons l'approche de Bathe et Sussman [5] qui nous mènera à la description des matériaux dit *semi-déformables*. Le point de départ est encore ici la relation contraintes-déformations que l'on suppose sous sa forme la plus générale, soit :

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\gamma}(\mathbf{u}) = C_{ijkl}\gamma_{kl}$$

◆ **Remarque 13.4**

Le tenseur d'élasticité d'ordre 4  $\mathbf{C}$  n'est pas tout-à-fait quelconque. On montre en effet que ce tenseur dérive généralement d'un potentiel  $\Phi$  suivant une expression de la forme :

$$C_{ijkl} = \frac{\partial^2 \Phi}{\partial \gamma_{ij} \partial \gamma_{kl}}$$

Puisque  $\gamma_{ij} = \gamma_{ji}$ , on en conclut immédiatement les propriétés de symétrie de ce tenseur :

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij}$$

Le tenseur  $\mathbf{C}$  possède donc au plus 21 composantes différentes et éventuellement non nulles. ◆

Nous souhaitons en arriver à une décomposition similaire à celle de la relation 13.4 pour les matériaux isotropes. Pour ce faire, on pose :

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\gamma}(\mathbf{u}) = (\mathbf{C} - \tilde{\mathbf{C}}) : \boldsymbol{\gamma}(\mathbf{u}) - p\tilde{\mathbf{I}} = \boldsymbol{\tau} - p\tilde{\mathbf{I}} \quad (13.9)$$

où  $\tilde{\mathbf{I}}$  et  $\tilde{\mathbf{C}}$  sont les tenseurs respectivement d'ordre 2 et 4 définis par les relations <sup>1</sup> :

$$\tilde{C}_{ijkl} = \frac{C_{aaij}C_{bbkl}}{C_{aabb}} \quad \text{et} \quad \tilde{I}_{ij} = \frac{3C_{aaij}}{C_{aabb}}$$

Pour que  $\boldsymbol{\sigma}$  soit toujours égal à  $\mathbf{C} : \boldsymbol{\gamma}(\mathbf{u})$ , la pression doit vérifier :

$$p\tilde{\mathbf{I}} = -\tilde{\mathbf{C}} : \boldsymbol{\gamma}(\mathbf{u}) \quad \text{ou encore} \quad p\tilde{I}_{ij} = -\tilde{C}_{ijkl}\gamma_{kl}$$

---

1. Rappelons que l'on utilise ici la notation d'Einstein sur les indices répétés.

qui peut encore s'écrire :

$$p \left( \frac{3\mathcal{C}_{aaij}}{\mathcal{C}_{aabb}} \right) = -\frac{\mathcal{C}_{aaij}\mathcal{C}_{bbkl}}{\mathcal{C}_{aabb}} \gamma_{kl}$$

ce qui entraîne que :

$$p = -\frac{\mathcal{C}_{aakl}\gamma_{kl}}{3}$$

ce qui devient enfin :

$$p = -\frac{\mathcal{C}_{aabb}}{9} \left( \frac{3\mathcal{C}_{aakl}}{\mathcal{C}_{aabb}} \right) \gamma_{kl} = -\left( \frac{\mathcal{C}_{aabb}}{9} \right) \tilde{\mathbf{I}} : \boldsymbol{\gamma}(\mathbf{u})$$

À l'aide de ces dernières expressions, on constate aisément que la relation 13.9 n'est qu'un réarrangement du tenseur  $\boldsymbol{\sigma}$  permettant d'introduire la pression. Plusieurs remarques s'imposent dès maintenant.

◆ **Remarque 13.5**

1.  $\mathcal{C}_{aabb} = \sum_{a=1}^3 \sum_{b=1}^3 \mathcal{C}_{aabb}$  est un scalaire et non un tenseur.
2. Puisque

$$\boldsymbol{\tau} = (\mathbf{C} - \tilde{\mathbf{C}}) : \boldsymbol{\gamma}(\mathbf{u}) \quad \text{ou encore} \quad \tau_{ij} = (\mathcal{C}_{ijkl} - \tilde{\mathcal{C}}_{ijkl})\gamma_{kl}$$

on a alors :

$$\text{tr}(\boldsymbol{\tau}) = \tau_{ii} = (\mathcal{C}_{iikl} - \tilde{\mathcal{C}}_{iikl})\gamma_{kl} = \left( \mathcal{C}_{iikl} - \frac{\mathcal{C}_{aaii}\mathcal{C}_{aakl}}{\mathcal{C}_{aabb}} \right) \gamma_{kl} = 0$$

et le tenseur  $\boldsymbol{\tau}$  est de trace nulle.

◆

◆ **Remarque 13.6**

Dans le cas isotrope, la relation contraintes-déformations prend la forme de l'équation 11.9 et on a :

1. La constante  $\mathcal{C}_{aabb}$  devient :

$$\mathcal{C}_{aabb} = \lambda I_{aa}I_{bb} + \mu(2I_{ab}I_{ab}) = 9\lambda + 6\mu$$

2. Le tenseur  $\tilde{\mathbf{I}}$  est égal au tenseur  $\mathbf{I}$ . En effet :

$$\tilde{I}_{ij} = \frac{3\mathcal{C}_{aaij}}{\mathcal{C}_{aabb}} = \frac{\lambda I_{aa}I_{ij} + 2\mu I_{ai}I_{aj}}{3\lambda + 2\mu} = \frac{3\lambda I_{ij} + 2\mu I_{ij}}{3\lambda + 2\mu} = I_{ij}$$

3. La pression devient :

$$\begin{aligned} p &= -\frac{\mathcal{C}_{aakl}\gamma_{kl}}{3} = -\frac{(\lambda I_{aa}I_{kl} + 2\mu I_{ak}I_{al})\gamma_{kl}}{3} \\ &= -\frac{3\lambda I_{kl}\gamma_{kl} + 2\mu I_{kl}\gamma_{kl}}{3} = -\frac{(3\lambda + 2\mu)\text{tr}(\boldsymbol{\gamma}(\mathbf{u}))}{3} \\ &= -k \nabla \cdot \mathbf{u} \end{aligned}$$



Le système mixte dans le cas anisotrope peut maintenant s'écrire :

$$\begin{cases} -\nabla \cdot [(\mathbf{C} - \tilde{\mathbf{C}}) : \boldsymbol{\gamma}(\mathbf{u}) - p\tilde{\mathbf{I}}] & = \mathbf{r} \\ -\frac{9}{\mathcal{C}_{aabb}}p - (\tilde{\mathbf{I}} : \boldsymbol{\gamma}(\mathbf{u})) & = 0 \end{cases}$$

qui est une généralisation de la forme variationnelle 13.6. Multipliant ces 2 équations par des fonctions tests  $\mathbf{w}$  et  $q$  et intégrant par parties, on trouve la formulation variationnelle :

$$\begin{aligned} \int_{\Omega} [(\mathbf{C} - \tilde{\mathbf{C}}) : \boldsymbol{\gamma}(\mathbf{u})] : \boldsymbol{\gamma}(\mathbf{w}) \, dv & - \int_{\Omega} p\tilde{\mathbf{I}} : \boldsymbol{\gamma}(\mathbf{w}) \, dv & = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv \\ - \int_{\Omega} (\tilde{\mathbf{I}} : \boldsymbol{\gamma}(\mathbf{u}))q \, dv & - \frac{9}{\mathcal{C}_{aabb}} \int_{\Omega} pq \, dv & = 0 \end{aligned}$$

et on retrouve encore la forme 13.6, mais cette fois avec les formes bilinéaires suivantes :

$$\begin{aligned} a(\mathbf{u}, \mathbf{w}) & = \int_{\Omega} [(\mathbf{C} - \tilde{\mathbf{C}}) : \boldsymbol{\gamma}(\mathbf{u})] : \boldsymbol{\gamma}(\mathbf{w}) \, dv \\ b(\mathbf{w}, q) & = - \int_{\Omega} (\tilde{\mathbf{I}} : \boldsymbol{\gamma}(\mathbf{w}))q \, dv \\ c(p, q) & = \frac{9}{\mathcal{C}_{aabb}} \int_{\Omega} pq \, dv \end{aligned}$$

◆ **Remarque 13.7**

Selon Bathe et Sussman [5], ce sont les matériaux pour lesquels la constante :

$$\frac{9}{\mathcal{C}_{aabb}}$$

tend vers l'infini qui requièrent l'emploi de méthodes mixtes. ◆

## 13.3 Résultats numériques

### 13.3.1 Plaque trouée en élongation

Nous avons déjà vu ce problème au chapitre 11 (voir la figure 11.5) et nous l'avons résolu par une méthode en déplacements seulement. Cette formulation avait alors montré ses faiblesse dans

le cas où  $\nu$  approchait 0,5 c.-à-d. dans le cas d'un matériau quasi-incompressible. Nous reprenons cette fois la résolution mais avec une méthode mixte. Nous utiliserons un élément quadrangulaire à 9 noeuds pour la vitesse et un élément à 4 noeuds pour la pression (voir la figure 6.4).

Les figures 13.1 à 13.3 reprennent celles du chapitre 11 en ajoutant au passage la pression (absente en formulation déplacements). Les déplacements de la figure 13.1 ont été amplifiés d'un facteur 3.

Contrairement à la formulation en déplacement seulement, la formulation mixte donne d'excellents résultats, même pour  $\nu = 0,4999999$  comme on peut le constater à la figure 13.4

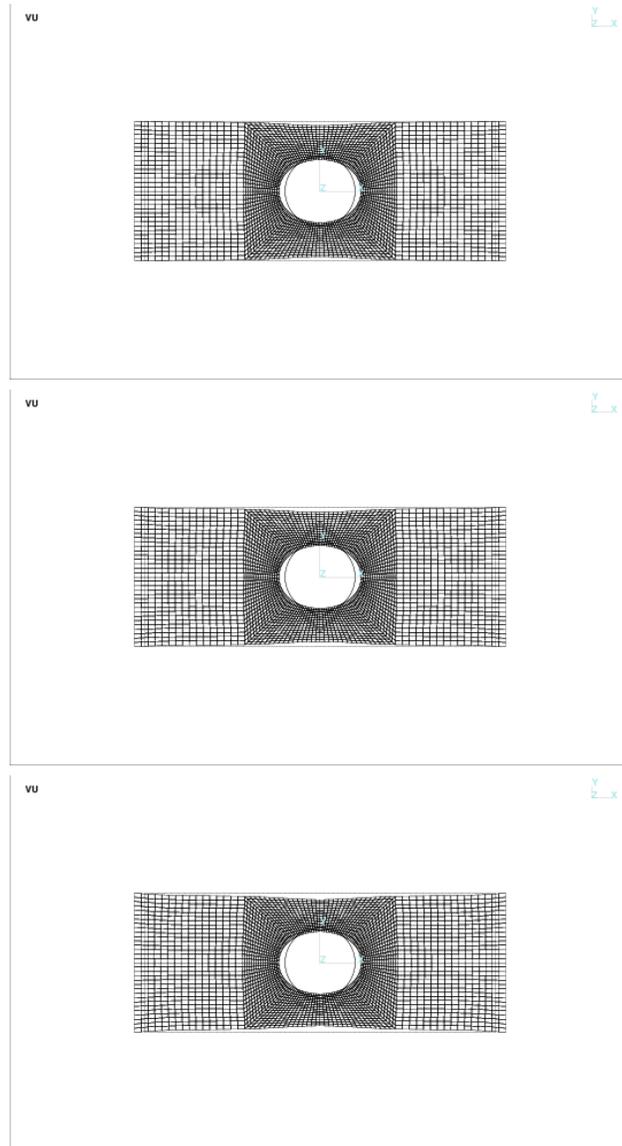


Figure 13.1 – Déplacements :  $\nu = 0,3$ ,  $\nu = 0,4$  et  $\nu = 0,4999$

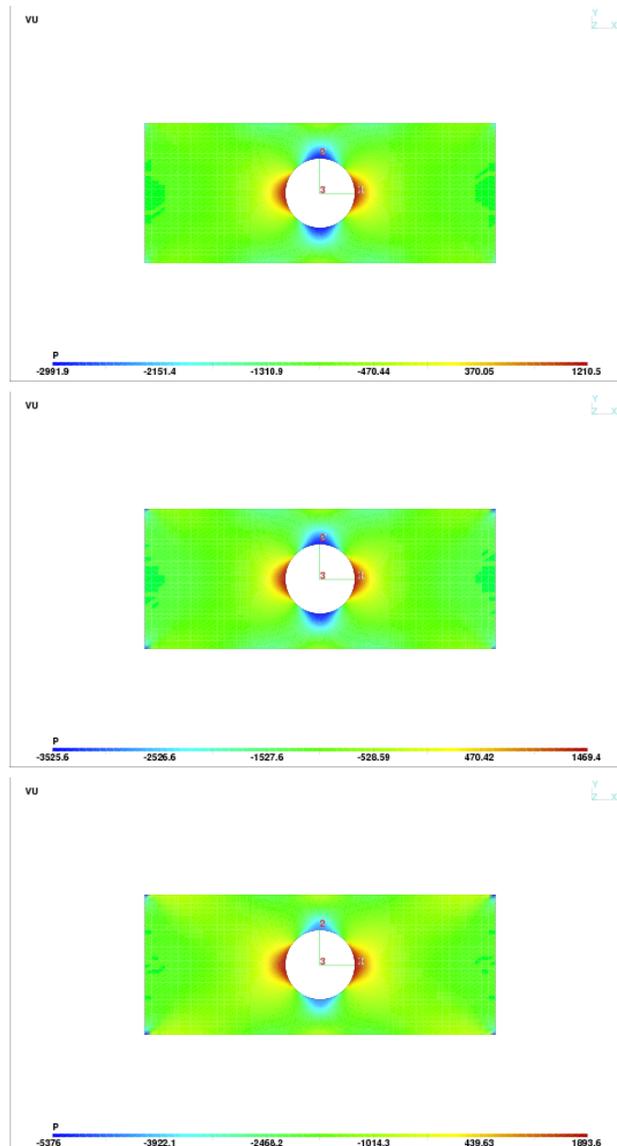


Figure 13.2 – Pression :  $\nu = 0,3$ ,  $\nu = 0,4$  et  $\nu = 0,4999$

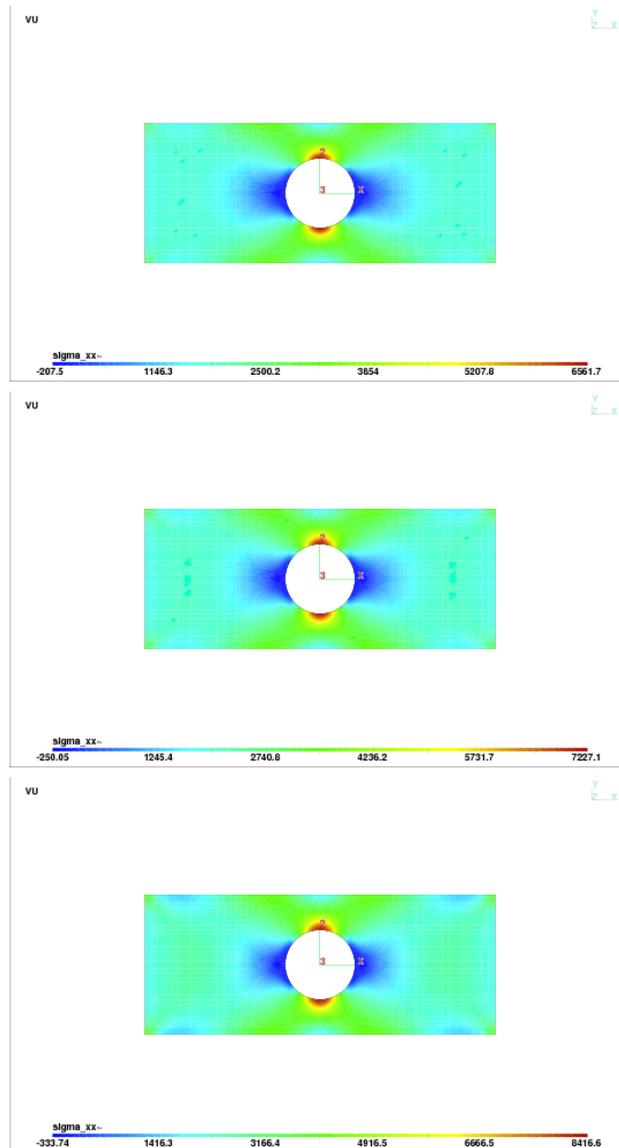


Figure 13.3 – Contrainte  $\sigma_{11}$  :  $\nu = 0,3$ ,  $\nu = 0,4$  et  $\nu = 0,4999$

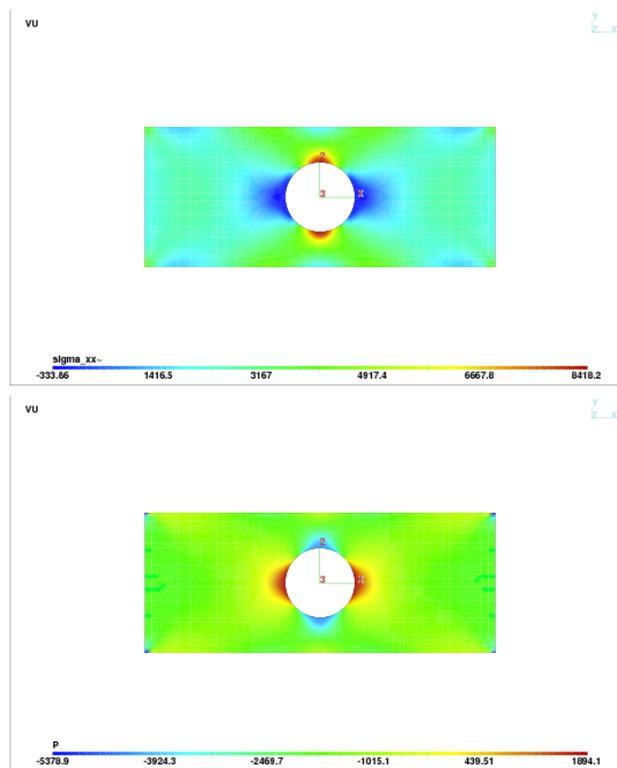


Figure 13.4 – Contrainte  $\sigma_{11}$  et  $p$  :  $\nu = 0,499\ 9999$

## 13.4 Exercices

1. La définition du tenseur des contraintes en élasticité linéaire du chapitre 11 nous donne :

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\gamma}(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I}$$

Poser  $p = -\lambda(\nabla \cdot \mathbf{u})$  et obtenir une nouvelle formulation mixte du problème d'élasticité linéaire (légèrement différente du système 13.6). Discuter de l'existence et de l'unicité de la solution.



# Chapitre 14

## Matériaux en grandes déformations

### 14.1 Introduction

Nous avons vu jusqu'à maintenant plusieurs applications de la méthode des éléments finis qui se caractérisaient toutes par un domaine de calcul fixe (noté  $\Omega$ ). En y regardant cependant de plus près, on découvre que ce n'était pas tout-à-fait le cas. En effet, les problèmes en élasticité linéaire rencontrés aux chapitres 11 et 13 entraînent une déformation du domaine  $\Omega$ . Cette déformation était supposée faible de sorte que, dans les formulations variationnelles, toutes les intégrales portaient quand même sur le domaine non déformé  $\Omega$ . Dans le cas des problèmes dits en grandes déformations, cette simplification n'est plus possible. Ces déformations entraînent des non-linéarités géométriques importantes. La normale à la géométrie n'est plus constante mais varie en fonction de la déformation. Par exemple, un pneu qui s'aplatit au contact de la route subit une déformation significative par rapport à sa forme circulaire initiale et on ne peut négliger l'évolution du domaine en fonction de la charge imposée. C'est l'objet de ce chapitre.

Le problème consiste à calculer la déformation d'un corps de forme initiale  $\Omega^0$  lorsqu'il est soumis à des déplacements et des contraintes externes. Nous ne considérerons pour le moment que les matériaux hyperélastiques. Au préalable, il nous faut introduire plusieurs concepts nouveaux qui nous permettront d'aborder les formulations variationnelles en grandes déformations.

### 14.2 Déformations et tenseurs associés

Nous noterons  $\Omega^0$  le domaine initial et  $\mathbf{X} = (X_1, X_2, X_3)^\top$  un point matériel de ce domaine. Au fil du temps, le domaine  $\Omega^0$  se déplace et se déforme et le domaine ainsi déformé sera noté  $\Omega^t$  (la notation indiquant la dépendance par rapport au temps). Le point correspondant à  $\mathbf{X}$  dans cette nouvelle configuration sera noté  $\mathbf{x}(\mathbf{X}, t) = (x_1(\mathbf{X}, t), x_2(\mathbf{X}, t), x_3(\mathbf{X}, t))^\top$ . L'indice supérieur  $\top$  dénote la transposée de sorte que les points et vecteurs sont représentés par des matrices colonnes 3 par 1. La particule initialement située en  $\mathbf{X}$  dans la configuration initiale  $\Omega^0$  sera donc située en  $\mathbf{x}$  dans la configuration  $\Omega^t$  au temps  $t$ . Le point matériel  $\mathbf{X}$  a donc subi un déplacement

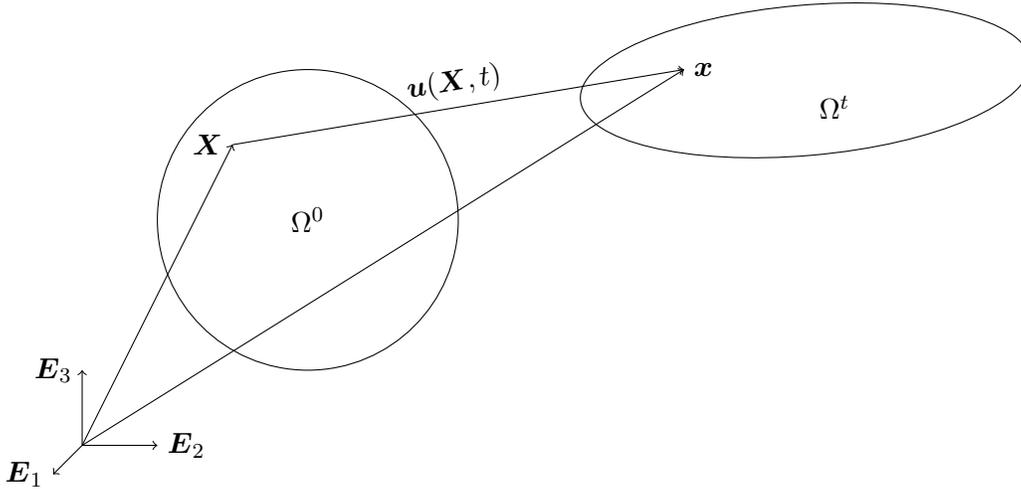


Figure 14.1 – Description cinématique

$\mathbf{u}(\mathbf{X}, t) = \mathbf{x}(\mathbf{X}, t) - \mathbf{X}$ , que l'on peut exprimer sous la forme :

$$\begin{bmatrix} x_1(\mathbf{X}, t) \\ x_2(\mathbf{X}, t) \\ x_3(\mathbf{X}, t) \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} u_1(\mathbf{X}, t) \\ u_2(\mathbf{X}, t) \\ u_3(\mathbf{X}, t) \end{bmatrix} \quad (14.1)$$

Des éléments de longueur  $d\mathbf{X}$ , de surface  $dA$  ou de volume  $dV$  dans la configuration initiale  $\Omega^0$  seront transformés respectivement en  $d\mathbf{x}$ ,  $da$  et  $dv$  dans la configuration déformée  $\Omega^t$ . La clé de ces transformations est le tenseur *gradient de déformation* défini par :

$$\mathbf{F} = \begin{bmatrix} \frac{\partial x_1}{\partial X_1} & \frac{\partial x_1}{\partial X_2} & \frac{\partial x_1}{\partial X_3} \\ \frac{\partial x_2}{\partial X_1} & \frac{\partial x_2}{\partial X_2} & \frac{\partial x_2}{\partial X_3} \\ \frac{\partial x_3}{\partial X_1} & \frac{\partial x_3}{\partial X_2} & \frac{\partial x_3}{\partial X_3} \end{bmatrix} \quad \text{dont l'inverse est } \mathbf{F}^{-1} = \begin{bmatrix} \frac{\partial X_1}{\partial x_1} & \frac{\partial X_1}{\partial x_2} & \frac{\partial X_1}{\partial x_3} \\ \frac{\partial X_2}{\partial x_1} & \frac{\partial X_2}{\partial x_2} & \frac{\partial X_2}{\partial x_3} \\ \frac{\partial X_3}{\partial x_1} & \frac{\partial X_3}{\partial x_2} & \frac{\partial X_3}{\partial x_3} \end{bmatrix} \quad (14.2)$$

Nous supposons que le tenseur gradient de déformation est toujours inversible et nous notons  $J$  son déterminant et  $F_{ij}$  ses composantes. En pratique, on calculera le déplacement  $\mathbf{u}$  de sorte qu'il convient d'utiliser la relation 14.1 et d'exprimer le tenseur gradient de déformation en fonction du tenseur gradient des déplacements :

$$\mathbf{F} = \mathbf{I} + \nabla_{\mathbf{X}} \mathbf{u} \quad (14.3)$$

L'indice inférieur dans le symbole  $\nabla_X$  indique que les dérivées sont prises par rapport aux variables  $X_i$  c.-à-d. dans la configuration initiale. On a ainsi :

$$(\nabla_X \mathbf{u})_{ij} = \frac{\partial u_i}{\partial X_j}$$

dont la partie symétrique :

$$\gamma_X(\mathbf{u}) = \frac{\nabla_X \mathbf{u} + \nabla_X^\top \mathbf{u}}{2}$$

est le tenseur de déformation que nous avons rencontré dans le cas des petites déformations (voir le chapitre 11). Par dérivation en chaîne, on a :

$$\frac{\partial u_i}{\partial x_j} = \frac{\partial u_i}{\partial X_k} \frac{\partial X_k}{\partial x_j}$$

de sorte que :

$$\nabla \mathbf{u} = \nabla_X \mathbf{u} \cdot \mathbf{F}^{-1} \quad \text{ou encore} \quad \nabla_X \mathbf{u} = \nabla \mathbf{u} \cdot \mathbf{F} \quad (14.4)$$

Il en résulte que le tenseur de déformation  $\gamma(\mathbf{u})$  peut s'écrire :

$$\gamma(\mathbf{u}) = \frac{1}{2} (\nabla \mathbf{u} + \nabla^\top \mathbf{u}) = \frac{1}{2} (\nabla_X \mathbf{u} \cdot \mathbf{F}^{-1} + \mathbf{F}^{-\top} \cdot \nabla_X^\top \mathbf{u}) \quad (14.5)$$

En multipliant la relation (14.3) de chaque côté par  $\mathbf{F}^{-1}$ , on a de plus que :

$$\mathbf{F}^{-1} = \mathbf{I} - \nabla \mathbf{u} \quad (14.6)$$

Le tenseur  $\mathbf{F}$  permet de suivre l'évolution de la déformation de la géométrie initiale. Ainsi, un élément de longueur  $d\mathbf{X} = (dX_1, dX_2, dX_3)^\top$  de la géométrie initiale sera transformé en un élément de longueur  $d\mathbf{x} = (dx_1, dx_2, dx_3)^\top$  sur la géométrie déformée. De plus, puisque :

$$dx_i = \frac{\partial x_i}{\partial X_j} dX_j$$

on a immédiatement que :

$$d\mathbf{x} = \mathbf{F} \cdot d\mathbf{X}$$

La longueur  $dl$  de l'élément  $d\mathbf{X}$  sera modifiée pour devenir :

$$(dl)^2 = d\mathbf{x} \cdot d\mathbf{x} = (\mathbf{F} \cdot d\mathbf{X}) \cdot (\mathbf{F} \cdot d\mathbf{X}) = d\mathbf{X} \cdot (\mathbf{F}^\top \cdot \mathbf{F} \cdot d\mathbf{X}) = d\mathbf{X} \cdot (\mathbf{C} \cdot d\mathbf{X}) = (\mathbf{C} \cdot d\mathbf{X}) \cdot d\mathbf{X}$$

par définition de la transposée d'un tenseur (voir l'annexe B pour quelques rappels sur les tenseurs) et où on a introduit le tenseur de Cauchy-Green  $\mathbf{C} = \mathbf{F}^\top \cdot \mathbf{F}$  qui jouera un rôle important par la suite.<sup>1</sup> Notons au passage que ce tenseur est symétrique, contrairement au tenseur gradient de déformation  $\mathbf{F}$ .

1. Le symbole « $\cdot$ » désigne le produit contracté de deux tenseurs ou le produit d'un tenseur et d'un vecteur.

De même, si on considère maintenant deux éléments de longueur  $d\mathbf{X}^1$  et  $d\mathbf{X}^2$  (non colinéaires), on peut définir un élément de surface  $d\mathbf{A} = d\mathbf{X}^1 \wedge d\mathbf{X}^2$  ( $\wedge$  étant le produit vectoriel). On a ainsi :

$$d\mathbf{A} = d\mathbf{X}^1 \wedge d\mathbf{X}^2 = \frac{d\mathbf{X}^1 \wedge d\mathbf{X}^2}{\|d\mathbf{X}^1 \wedge d\mathbf{X}^2\|} \|d\mathbf{X}^1 \wedge d\mathbf{X}^2\| = \mathbf{N}dA$$

en rappelant que le produit vectoriel donne un vecteur orthogonal à  $d\mathbf{X}^1$  et  $d\mathbf{X}^2$  et que le module du produit vectoriel n'est autre que l'aire du parallélogramme engendré par ces deux vecteurs. On remarquera de plus la différence entre l'élément d'aire orienté  $d\mathbf{A}$  (un vecteur) et son aire  $dA$  (un scalaire).

♣ **Lemme 14.4 (Changement de volume) :**

Un élément de volume  $dV$  de la configuration initiale sera transformé en un élément de volume  $dv$  de la configuration déformée suivant la formule classique de changement de variables :

$$dv = \det(\mathbf{F}) dV = J dV \quad \text{ou encore} \quad d\mathbf{x}^1 \cdot (d\mathbf{x}^2 \wedge d\mathbf{x}^3) = J d\mathbf{X}^1 \cdot (d\mathbf{X}^2 \wedge d\mathbf{X}^3) \quad (14.7)$$

**Démonstration :**

Rappelons tout d'abord un résultat classique dont la démonstration est laissée en exercice. Si  $\mathbf{M}$  est un tenseur et  $\mathbf{u}, \mathbf{v}$  deux vecteurs, alors :

$$(\mathbf{M}\mathbf{u} \wedge \mathbf{M}\mathbf{v}) = \det(\mathbf{M}) \mathbf{M}^{-\top} \cdot (\mathbf{u} \wedge \mathbf{v})$$

On a ainsi :

$$\begin{aligned} dv &= d\mathbf{x}^1 \cdot (d\mathbf{x}^2 \wedge d\mathbf{x}^3) = \mathbf{F} \cdot d\mathbf{X}^1 \cdot (\mathbf{F} \cdot d\mathbf{X}^2 \wedge \mathbf{F} \cdot d\mathbf{X}^3) \\ &= \det(\mathbf{F}) \mathbf{F} \cdot d\mathbf{X}^1 \cdot (\mathbf{F}^{-\top} \cdot (d\mathbf{X}^2 \wedge d\mathbf{X}^3)) \\ &= \det(\mathbf{F}) d\mathbf{X}^1 \cdot (d\mathbf{X}^2 \wedge d\mathbf{X}^3) = JdV \end{aligned}$$

♣

De même, un élément d'aire orientée  $d\mathbf{A} = \mathbf{N}dA$  et sa normale  $\mathbf{N}$  sur la géométrie initiale (non déformée) seront transformés respectivement en  $d\mathbf{a} = \mathbf{n}da$  et  $\mathbf{n}$  sur la géométrie actuelle (déformée) selon la formule de Nanson que nous allons démontrer.

♣ **Lemme 14.5 (Formule de Nanson) :**

Un élément d'aire orientée se transforme suivant la relation :

$$d\mathbf{a} = J(\mathbf{X})\mathbf{F}^{-\top} \cdot d\mathbf{A} \quad \text{ou encore} \quad \mathbf{n}da = J(\mathbf{X})\mathbf{F}^{-\top} \cdot \mathbf{N}dA \quad (14.8)$$

**Démonstration :**

Puisque  $dv = J dV$ , on a :

$$d\mathbf{x} \cdot d\mathbf{a} = J d\mathbf{X} \cdot d\mathbf{A}$$

et puisque  $d\mathbf{x} = \mathbf{F} \cdot d\mathbf{X}$  :

$$(\mathbf{F} \cdot d\mathbf{X}) \cdot d\mathbf{a} = J d\mathbf{X} \cdot d\mathbf{A} \text{ ou encore } d\mathbf{X} \cdot (\mathbf{F}^\top \cdot d\mathbf{a}) = J d\mathbf{X} \cdot d\mathbf{A}$$

pour tout vecteur  $d\mathbf{X}$ . On a ainsi :

$$\mathbf{F}^\top \cdot d\mathbf{a} = J d\mathbf{A}$$

d'où le résultat. ♣

♣ **Lemme 14.6 (Formule de Piola) :**

L'opérateur de divergence se transforme d'une géométrie à l'autre suivant la relation :

$$\int_{\Omega^t} \nabla \cdot \mathbf{w} dv = \int_{\Omega^0} \nabla_X \cdot \mathbf{W} dV \text{ où } \mathbf{W} = J\mathbf{F}^{-1} \cdot \mathbf{w} \quad (14.9)$$

**Démonstration :**

Le résultat découle de la formule de Nanson et du théorème de la divergence. On a en effet :

$$\int_{\Omega^t} \nabla \cdot \mathbf{w} dv = \int_{\Gamma^t} \mathbf{w} \cdot \mathbf{n} da = \int_{\Gamma^0} \mathbf{w} \cdot (J\mathbf{F}^{-\top} \cdot \mathbf{N}) dA = \int_{\Gamma^0} J(\mathbf{F}^{-1} \cdot \mathbf{w}) \cdot \mathbf{N} dA = \int_{\Gamma^0} \mathbf{W} \cdot \mathbf{N} dA$$

et on obtient le résultat par une nouvelle application du théorème de la divergence, cette fois sur la géométrie initiale. ♣

♣ **Lemme 14.7 (Jacobien surfacique et évolution de la normale) :**

Le rapport des aires entre des éléments de surface déformée et non déformée  $da/dA$  est appelé le *jacobien surfacique* que l'on note  $J_s$  et qui s'écrit :

$$J_s = \frac{da}{dA} = J \|\mathbf{F}^{-\top} \cdot \mathbf{N}\| = J \sqrt{(\mathbf{C}^{-1} \cdot \mathbf{N}) \cdot \mathbf{N}} = \frac{J}{\sqrt{(\mathbf{F} \cdot \mathbf{F}^\top \cdot \mathbf{n}) \cdot \mathbf{n}}} \quad (14.10)$$

De plus, la normale à la configuration déformée  $\mathbf{n}$  évolue à partir de la normale à la configuration initiale  $\mathbf{N}$  suivant la formule :

$$\mathbf{n} = J\mathbf{F}^{-\top} \cdot \mathbf{N} \frac{dA}{da} = \frac{J}{J_s} (\mathbf{F}^{-\top} \cdot \mathbf{N}) \quad (14.11)$$

**Démonstration :**

De la formule de Nanson 14.8, on a  $\mathbf{n} da = J \mathbf{F}^{-\top} \cdot \mathbf{N} dA$  de sorte qu'en prenant la norme de chaque côté, on a d'une part :

$$da = J \|\mathbf{F}^{-\top} \cdot \mathbf{N}\| dA \quad \text{ou encore} \quad \frac{da}{dA} = J \|\mathbf{F}^{-\top} \cdot \mathbf{N}\|$$

et d'autre part :

$$\|\mathbf{F}^{-\top} \cdot \mathbf{N}\|^2 = (\mathbf{F}^{-\top} \cdot \mathbf{N}) \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) = (\mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} \cdot \mathbf{N}) \cdot \mathbf{N} = (\mathbf{C}^{-1} \cdot \mathbf{N}) \cdot \mathbf{N}$$

d'où :

$$\frac{da}{dA} = J \sqrt{(\mathbf{C}^{-1} \cdot \mathbf{N}) \cdot \mathbf{N}} \quad \text{et par un calcul similaire} \quad \frac{da}{dA} = \frac{J}{\sqrt{(\mathbf{F} \cdot \mathbf{F}^{\top} \cdot \mathbf{n}) \cdot \mathbf{n}}}$$



♣ **Lemme 14.8 (Évolution de la masse volumique) :**

Puisqu'il n'y a aucune perte de masse au cours de la déformation du corps, la masse volumique  $\rho$  évolue suivant la formule :

$$\rho_0 dV = \rho dv = \rho J dV \quad \text{ou encore} \quad \rho = \frac{\rho_0}{J} \quad (14.12)$$

de sorte que si dét  $\mathbf{F} = J = 1$ , la masse volumique reste constante. ♣

★ **Théorème 14.1 (Théorème de transport de Reynolds)**

Soit  $\Omega^t$  un domaine se déplaçant dans le temps à une vitesse  $\mathbf{v}$  et  $f$  une fonction régulière. On a alors

$$\frac{d}{dt} \int_{\Omega^t} f(\mathbf{x}, t) dv = \int_{\Omega^t} \frac{\partial f}{\partial t} + \nabla \cdot (f\mathbf{v}) dv = \int_{\Omega^t} \frac{\partial f}{\partial t} dv + \int_{\partial\Omega^t} f\mathbf{v} \cdot \mathbf{n} ds \quad (14.13)$$

**Démonstration :** La difficulté provient du domaine d'intégration qui varie dans le temps. On

contourne cette difficulté en se ramenant sur le domaine de référence  $\Omega^0$  qui lui est constant.

$$\begin{aligned}
\frac{d}{dt} \int_{\Omega^t} f(\mathbf{x}, t) dv &= \frac{d}{dt} \int_{\Omega^0} f(\mathbf{x}, t) J dV = \int_{\Omega^0} \frac{d}{dt} (f(\mathbf{x}, t) J(\mathbf{X}, t)) dV \\
&= \int_{\Omega^0} J \frac{d}{dt} f(\mathbf{x}, t) + f \frac{d}{dt} J dV \\
&= \int_{\Omega^0} J \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial t} \right) + f \frac{d}{dt} J dV \\
&= \int_{\Omega^0} \left( \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f \right) J + f J \nabla \cdot \mathbf{v} dV \\
&= \int_{\Omega^0} \left( \frac{\partial f}{\partial t} + \nabla \cdot (f \mathbf{v}) \right) J dV \\
&= \int_{\Omega^t} \left( \frac{\partial f}{\partial t} + \nabla \cdot (f \mathbf{v}) \right) dv
\end{aligned}$$

et la dernière égalité de 14.13 provient du théorème de la divergence A.1. ★

### 14.3 Tenseurs de Green-Lagrange et Piola-Kirchoff

Les lois de comportement du matériau reliant les contraintes aux déformations sont d'une grande complexité. Nous introduirons dans ce qui suit les tenseurs de Green-Lagrange et de Piola-Kirchoff apparaissant dans ces lois de comportement. Nous discutons d'abord d'un résultat de décomposition des tenseurs d'ordre 2.

#### ★ Théorème 14.2 (Décomposition polaire)

Tout tenseur peut s'écrire comme le produit d'un tenseur de rotation  $\mathbf{R}$  orthogonal ( $\mathbf{R}^\top \mathbf{R} = \mathbf{R} \mathbf{R}^\top = \mathbf{I}$ ) et d'un tenseur d'élongation  $\mathbf{U}$  (symétrique et défini positif) ou, dans l'ordre inverse, d'un tenseur d'élongation symétrique et défini positif  $\mathbf{V}$  et d'un tenseur de rotation  $\mathbf{R}$ . En particulier, le tenseur de déformation  $\mathbf{F}$  peut être factorisé sous la forme :

$$\mathbf{F} = \mathbf{R} \cdot \mathbf{U} = \mathbf{V} \cdot \mathbf{R}$$

**Démonstration** (facultative) :

Puisque  $\mathbf{F}$  est inversible, le tenseur  $\mathbf{F}^\top \cdot \mathbf{F}$  est symétrique et défini positif de sorte qu'il peut être diagonalisé. On a donc :

$$\mathbf{F}^\top \cdot \mathbf{F} = \mathbf{Q} \cdot \Lambda^2 \cdot \mathbf{Q}^\top \quad (14.14)$$

où  $\mathbf{Q}$  est un tenseur orthogonal ( $\mathbf{Q}^\top \cdot \mathbf{Q} = \mathbf{Q} \cdot \mathbf{Q}^\top = \mathbf{I}$ ),  $\mathbf{\Lambda}$  est un tenseur diagonal contenant la racine carrée des valeurs propres de  $\mathbf{F}^\top \cdot \mathbf{F}$  (puisqu'elles sont toutes strictement positives). On pose maintenant :

$$\mathbf{U} = \mathbf{Q} \cdot \mathbf{\Lambda} \cdot \mathbf{Q}^\top \quad \text{d'où} \quad \mathbf{U}^2 = (\mathbf{Q} \cdot \mathbf{\Lambda} \cdot \mathbf{Q}^\top)(\mathbf{Q} \cdot \mathbf{\Lambda} \cdot \mathbf{Q}^\top) = \mathbf{Q} \cdot \mathbf{\Lambda}^2 \cdot \mathbf{Q}^\top = \mathbf{F}^\top \cdot \mathbf{F} = \mathbf{C}$$

Le tenseur  $\mathbf{U}$  est de toute évidence symétrique et défini positif et est en quelque sorte la racine carrée de  $\mathbf{F}^\top \cdot \mathbf{F}$ . Posant maintenant  $\mathbf{R} = \mathbf{F} \cdot \mathbf{U}^{-1}$ , on constate que ce nouveau tenseur est orthogonal. En effet :

$$\mathbf{R}^\top \cdot \mathbf{R} = \mathbf{U}^{-\top} \cdot (\mathbf{F}^\top \cdot \mathbf{F}) \cdot \mathbf{U}^{-1} = (\mathbf{Q}^{-\top} \cdot \mathbf{\Lambda}^{-1} \cdot \mathbf{Q}^{-1}) \cdot (\mathbf{Q} \cdot \mathbf{\Lambda}^2 \cdot \mathbf{Q}^\top) \cdot (\mathbf{Q}^{-\top} \cdot \mathbf{\Lambda}^{-1} \cdot \mathbf{Q}^{-1}) = \mathbf{U}^{-\top} \cdot \mathbf{U}^2 \cdot \mathbf{U}^{-1} = \mathbf{I}$$

puisque  $\mathbf{U}$  est symétrique. Suivant ces définitions, on a bien  $\mathbf{F} = \mathbf{R} \cdot \mathbf{U}$  et en définissant le tenseur symétrique  $\mathbf{V} = \mathbf{R} \cdot \mathbf{U} \cdot \mathbf{R}^\top$ , on a bien  $\mathbf{V} \cdot \mathbf{R} = \mathbf{R} \cdot \mathbf{U} = \mathbf{F}$ . De plus,

$$\mathbf{V} = \mathbf{R} \cdot \mathbf{U} \cdot \mathbf{R}^\top = \mathbf{R} \cdot \mathbf{Q} \cdot \mathbf{\Lambda} \cdot \mathbf{Q}^\top \cdot \mathbf{R}^\top = (\mathbf{R} \cdot \mathbf{Q}) \cdot \mathbf{\Lambda} \cdot (\mathbf{R} \cdot \mathbf{Q})^\top$$

ce qui permet de conclure que  $\mathbf{U}$  et  $\mathbf{V}$  possèdent les mêmes valeurs propres mais que les vecteurs propres de  $\mathbf{V}$  sont obtenus en effectuant une rotation des vecteurs propres de  $\mathbf{U}$ . ★

#### Définition 14.1

Le tenseur de déformation de Green-Lagrange, noté  $\mathbf{E}$ , est défini par :

$$\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I}) = \frac{1}{2}(\mathbf{U}^2 - \mathbf{I})$$

et le tenseur de Green-Lagrange est donc insensible aux rotations puisqu'il ne dépend pas de  $\mathbf{R}$  (contrairement au tenseur de déformation  $\mathbf{F}$ ).

On peut de plus exprimer le tenseur de Green-Lagrange en fonction des composantes du déplacement. En effet, on a en vertu de la relation 14.3 :

$$\mathbf{E} = \frac{1}{2}(\mathbf{F}^\top \cdot \mathbf{F} - \mathbf{I}) = \frac{1}{2} \left( (\mathbf{I} + \nabla_X \mathbf{u})^\top \cdot (\mathbf{I} + \nabla_X \mathbf{u}) - \mathbf{I} \right)$$

et on obtient immédiatement l'expression :

$$\mathbf{E} = \frac{1}{2} \left( \nabla_X \mathbf{u} + \nabla_X^\top \mathbf{u} + \nabla_X^\top \mathbf{u} \cdot \nabla_X \mathbf{u} \right) = \gamma_X(\mathbf{u}) + \frac{1}{2} \nabla_X^\top \mathbf{u} \cdot \nabla_X \mathbf{u} \quad (14.15)$$

#### ◆ Remarque 14.1

Lorsque les déformations sont petites, on peut négliger les termes d'ordre 2 dans le tenseur de Green-Lagrange de sorte que :

$$\mathbf{E} \simeq \frac{1}{2} \left( \nabla_X \mathbf{u} + \nabla_X^\top \mathbf{u} \right) = \gamma_X(\mathbf{u})$$

C'est cette expression qui a été utilisée en élasticité linéaire. D'autre part, l'expression 14.15 est bien une égalité, en ce sens qu'aucun terme n'a été négligé. ♦

Le tenseur le plus important est sans doute le tenseur des contraintes  $\boldsymbol{\sigma}$ , qui agit sur la configuration déformée  $\Omega^t$ . C'est donc ce tenseur que nous essaierons d'évaluer. Il est cependant difficile de le manipuler directement puisque la géométrie varie constamment. Il est donc souvent préférable de revenir sur la configuration initiale  $\Omega^0$  et y définir des tenseurs *de passage*, qui nous permettront éventuellement de calculer  $\boldsymbol{\sigma}$ .

**Définition 14.2**

Si  $d\mathbf{f}$  désigne un élément de force agissant sur un élément de surface  $da$  de normale  $\mathbf{n}$  de  $\Omega^t$ , le tenseur de Cauchy  $\boldsymbol{\sigma}$  est celui qui vérifie :

$$d\mathbf{f} = \boldsymbol{\sigma} \cdot \mathbf{n} da \quad \text{ou encore} \quad \frac{d\mathbf{f}}{da} = \boldsymbol{\sigma} \cdot \mathbf{n} \quad (14.16)$$

**Définition 14.3**

Si  $d\mathbf{f}$  désigne un élément de force agissant sur un élément de surface  $da$  de normale  $\mathbf{n}$  de  $\Omega^t$ , alors on définit le premier tenseur de Piola-Kirchoff  $\mathbf{\Pi}$  par la relation :

$$d\mathbf{f} = \mathbf{\Pi} \cdot \mathbf{N} dA \quad (14.17)$$

que l'on peut aussi écrire :

$$\frac{d\mathbf{f}}{dA} = \mathbf{\Pi} \cdot \mathbf{N}$$

de sorte que le premier tenseur de Piola-Kirchoff exprime une force sur la configuration actuelle par unité d'aire non déformée.

**Définition 14.4**

Si  $d\mathbf{f} = \boldsymbol{\sigma} \cdot \mathbf{n}$  désigne un élément de force agissant sur un élément de surface  $da$  de normale  $\mathbf{n}$  de  $\Omega^t$ , alors, on définit sur la configuration initiale  $\Omega^0$ , le deuxième tenseur de Piola-Kirchoff  $\mathbf{S}$  comme le tenseur vérifiant :

$$\mathbf{S} \cdot \mathbf{N} dA = \mathbf{F}^{-1} \cdot d\mathbf{f} = d\mathbf{f}_0 \quad (14.18)$$

Il s'agit donc d'une force sur la configuration initiale par unité d'aire non déformée.

Une comparaison des relations 14.16 et 14.18 montrent bien la similarité des rôles de  $\boldsymbol{\sigma}$  sur  $\Omega^t$  et de  $\mathbf{S}$  sur  $\Omega^0$ .

**♣ Lemme 14.9**

Les premier et deuxième tenseurs de Piola-Kirchoff  $\mathbf{\Pi}$  et  $\mathbf{S}$  sont reliés par la relation :

$$\mathbf{\Pi} = \mathbf{F} \cdot \mathbf{S} \quad (14.19)$$

**Démonstration :**

On a d'une part :

$$d\mathbf{f} = \mathbf{\Pi} \cdot \mathbf{N} dA$$

et d'autre part :

$$d\mathbf{f} = \mathbf{F} \cdot d\mathbf{f}_0 = \mathbf{F} \cdot \mathbf{S} \cdot \mathbf{N} dA$$

En soustrayant les deux dernières expressions, on trouve le résultat puisque  $\mathbf{N} dA$  est arbitraire. ♣

Il nous faut également relier entre eux le tenseur des contraintes  $\boldsymbol{\sigma}$  sur la géométrie actuelle  $\Omega^t$ , au deuxième tenseur de Piola-Kirchoff. C'est l'objet du prochain résultat.

**♣ Lemme 14.10**

Le passage entre le tenseur de contraintes  $\boldsymbol{\sigma}$  et le deuxième tenseur de Piola-Kirchoff s'effectue à travers la relation :

$$\boldsymbol{\sigma} = \frac{1}{J} \mathbf{F} \cdot \mathbf{S} \cdot \mathbf{F}^\top \quad \text{ou encore} \quad \mathbf{S} = J \mathbf{F}^{-1} \cdot \boldsymbol{\sigma} \cdot \mathbf{F}^{-\top} \quad (14.20)$$

**Démonstration :**

Par définition du deuxième tenseur de Piola-Kirchoff, on a d'une part :

$$d\mathbf{f} = \mathbf{F} \cdot d\mathbf{f}_0 = \mathbf{F} \cdot \mathbf{S} \cdot (\mathbf{N} dA)$$

et d'autre part, par la formule de Nanson, on a :

$$d\mathbf{f} = \boldsymbol{\sigma} \cdot \mathbf{n} da = \boldsymbol{\sigma} \cdot \left( J \mathbf{F}^{-\top} \cdot (\mathbf{N} dA) \right)$$

Soustrayant ces deux dernières relations, on trouve :

$$\left( J\boldsymbol{\sigma} \cdot \mathbf{F}^{-\top} - \mathbf{F} \cdot \mathbf{S} \right) \cdot (\mathbf{N} dA) = 0$$

d'où le résultat puisqu'ici encore  $\mathbf{N} dA$  est arbitraire. ♣

#### ♣ Lemme 14.11

Le tenseur de Cauchy et le premier tenseur de Piola-Kirchoff sont liés par les relations :

$$\boldsymbol{\sigma} = \frac{1}{J} \boldsymbol{\Pi} \cdot \mathbf{F}^{\top} \quad \text{ou encore} \quad \boldsymbol{\Pi} = J\boldsymbol{\sigma} \cdot \mathbf{F}^{-\top} \quad (14.21)$$

**Démonstration** : immédiate ♣

## 14.4 Matériaux hyperélastiques

Les matériaux hyperélastiques couvrent une vaste gamme d'applications notamment dans le domaine du pneumatique. Nous allons procéder dans un cadre assez général et nous verrons, dans un premier temps, les modèles de base (Saint-Venant-Kirchoff, néo-hookéen) et nous nous attarderons par la suite aux matériaux hyperélastiques incompressibles.

Pour un matériau hyperélastique, le deuxième tenseur de Piola-Kirchoff découle d'un potentiel d'énergie de déformation comme suit :

$$\mathbf{S} = \frac{\partial \Psi}{\partial \mathbf{E}} = 2 \frac{\partial \Psi}{\partial \mathbf{C}}$$

où  $\mathbf{C}$  est le tenseur de Cauchy-Green. Le potentiel  $\Psi$  dépend donc de  $\mathbf{C}$  (et donc de  $\mathbf{E}$ ) et nous l'exprimerons en fonction de ses invariants (ou de ceux de  $\mathbf{E}$ ) :

$$\begin{aligned} I_1 &= \text{tr}(\mathbf{C}) = C_{kk} \\ I_2 &= \frac{1}{2} (I_1^2 - \text{tr}(\mathbf{C} \cdot \mathbf{C})) = \frac{1}{2} (I_1^2 - (\mathbf{C} : \mathbf{C})) \\ I_3 &= \det \mathbf{C} = \det \mathbf{F}^{\top} \cdot \mathbf{F} = J^2 \end{aligned}$$

On remarque que  $\mathbf{C} : \mathbf{C} = I_1^2 - 2I_2$  est aussi un invariant. On supposera donc une relation de la forme :

$$\Psi = \Psi(I_1, I_2, I_3)$$

Le potentiel élastique exprimé ainsi est indépendant de la base utilisée. Le tenseur de Piola-Kirchoff devient donc :

$$\mathbf{S} = \left( \frac{\partial \Psi}{\partial I_1} \frac{\partial I_1}{\partial \mathbf{E}} + \frac{\partial \Psi}{\partial I_2} \frac{\partial I_2}{\partial \mathbf{E}} + \frac{\partial \Psi}{\partial I_3} \frac{\partial I_3}{\partial \mathbf{E}} \right) = 2 \left( \frac{\partial \Psi}{\partial I_1} \frac{\partial I_1}{\partial \mathbf{C}} + \frac{\partial \Psi}{\partial I_2} \frac{\partial I_2}{\partial \mathbf{C}} + \frac{\partial \Psi}{\partial I_3} \frac{\partial I_3}{\partial \mathbf{C}} \right)$$

relation qui fait intervenir les dérivées des invariants dont on trouvera les expressions à l'équation B.18 de l'annexe B.

### Modèle de Saint-Venant-Kirchoff

Le modèle de Saint-Venant-Kirchoff est la généralisation aux grandes déformations du modèle utilisé en élasticité linéaire (petites déformations). Le potentiel d'énergie prend la forme (voir par exemple Bonet et Wood [6]) :

$$\Psi = \frac{1}{2} \lambda (\text{tr} \mathbf{E})^2 + \mu (\mathbf{E} : \mathbf{E})$$

de sorte que :

$$\mathbf{S} = \frac{\partial \Psi}{\partial \mathbf{E}} = \lambda (\text{tr} \mathbf{E}) \mathbf{I} + 2\mu \mathbf{E}$$

Au moment de la linéarisation par la méthode de Newton, nous aurons besoin du tenseur d'élasticité d'ordre 4 :

$$\mathbf{C} = \frac{\partial^2 \Psi}{\partial \mathbf{E} \partial \mathbf{E}} = \frac{\partial \mathbf{S}}{\partial \mathbf{E}}$$

On a alors en utilisant les dérivées secondes des invariants (voir l'équation B.19) :

$$\frac{\partial \mathbf{S}}{\partial \mathbf{E}} : \boldsymbol{\delta}_E = \lambda \left( \frac{\partial \text{tr}(\mathbf{E})}{\partial \mathbf{E}} : \boldsymbol{\delta}_E \right) \mathbf{I} + 2\mu \frac{\partial \mathbf{E}}{\partial \mathbf{E}} : \boldsymbol{\delta}_E = \lambda (\mathbf{I} : \boldsymbol{\delta}_E) \mathbf{I} + 2\mu \boldsymbol{\mathcal{I}} : \boldsymbol{\delta}_E = (\lambda \mathbf{I} \otimes \mathbf{I} + 2\mu \boldsymbol{\mathcal{I}}) : \boldsymbol{\delta}_E$$

Le tenseur d'ordre 4  $\mathbf{C}$  a donc pour composantes :

$$C_{ijkl} = \lambda I_{ij} I_{kl} + 2\mu \left( \frac{I_{ik} I_{jl} + I_{il} I_{jk}}{2} \right)$$

résultat que nous avons déjà obtenu au chapitre 11 et qui montre bien que le modèle de Saint-Venant-Kirchoff est une généralisation du modèle d'élasticité linéaire<sup>2</sup>.

### Modèle néo-hookéen (compressible)

Un autre modèle hyperélastique est le modèle néo-hookéen pour lequel le potentiel est de la forme (voir [6]) :

$$\Psi = \frac{\mu}{2} (I_1 - 3) + \frac{\lambda}{2} (\ln J)^2 - \mu \ln J = \frac{\mu}{2} (I_1 - 3) + \frac{\lambda}{8} (\ln I_3)^2 - \frac{\mu}{2} \ln I_3$$

2. Voir l'annexe B pour la définition du tenseur  $\boldsymbol{\mathcal{I}}$  d'ordre 4.

de sorte que :

$$\begin{aligned}
\mathbf{S} &= 2 \frac{\partial \Psi}{\partial \mathbf{C}} = 2 \left( \frac{\mu}{2} \frac{\partial I_1}{\partial \mathbf{C}} + \frac{(\lambda \ln I_3)}{4I_3} \frac{\partial I_3}{\partial \mathbf{C}} - \frac{\mu}{2I_3} \frac{\partial I_3}{\partial \mathbf{C}} \right) \\
&= 2 \left[ \frac{\mu}{2} \mathbf{I} + \left( \frac{\lambda \ln(I_3)}{4} \mathbf{C}^{-1} - \frac{\mu}{2} \mathbf{C}^{-1} \right) \right] \quad \left( \text{puisque } \frac{\partial I_3}{\partial \mathbf{C}} = I_3 \mathbf{C}^{-1} \right) \\
&= \mu(\mathbf{I} - \mathbf{C}^{-1}) + \frac{\lambda}{2} (\ln I_3) \mathbf{C}^{-1}
\end{aligned}$$

et :

$$\begin{aligned}
\mathbf{C} : \delta \mathbf{C} &= 2 \frac{\partial \mathbf{S}}{\partial \mathbf{C}} : \delta \mathbf{C} \\
&= 2 \left\{ -\mu \frac{\partial \mathbf{C}^{-1}}{\partial \mathbf{C}} : \delta \mathbf{C} + \frac{\lambda}{2} \left[ \left( \frac{\partial \ln(I_3)}{\partial \mathbf{C}} : \delta \mathbf{C} \right) \mathbf{C}^{-1} + \ln(I_3) \left( \frac{\partial \mathbf{C}^{-1}}{\partial \mathbf{C}} : \delta \mathbf{C} \right) \right] \right\} \\
&= -2\mu \mathcal{J} : \delta \mathbf{C} + \lambda \left[ \left( \frac{1}{I_3} I_3 \mathbf{C}^{-1} : \delta \mathbf{C} \right) \mathbf{C}^{-1} + \ln(I_3) \mathcal{J} : \delta \mathbf{C} \right] \\
&= -2\mu \mathcal{J} : \delta \mathbf{C} + \lambda (\mathbf{C}^{-1} : \delta \mathbf{C}) \mathbf{C}^{-1} + \lambda \ln(I_3) \mathcal{J} : \delta \mathbf{C} \\
&= [(\lambda \ln(I_3) - 2\mu) \mathcal{J} + \lambda \mathbf{C}^{-1} \otimes \mathbf{C}^{-1}] : \delta \mathbf{C}
\end{aligned}$$

(Voir l'annexe B pour la définition du tenseur  $\mathcal{J}$  d'ordre 4).

#### ◆ Remarque 14.2

Si on considère le cas des petites déformations, on peut considérer le domaine comme constant et il n'y a plus de différences entre la variable  $\mathbf{X}$  sur la géométrie initiale et la variable  $\mathbf{x}$  de la géométrie déformée ( $\nabla_{\mathbf{X}} = \nabla$ ), ces deux géométries étant confondues. On peut aussi négliger les termes d'ordre 2 du tenseur de Cauchy-Green et on a :

$$\mathbf{C} = (\mathbf{F}^\top \cdot \mathbf{F}) = (\mathbf{I} + \nabla \mathbf{u})^\top \cdot (\mathbf{I} + \nabla \mathbf{u}) \simeq \mathbf{I} + 2\gamma(\mathbf{u})$$

de sorte que :

$$I_1 = \text{tr}(\mathbf{C}) \simeq \text{tr}(\mathbf{I} + 2\gamma(\mathbf{u})) = 3 + 2\nabla \cdot \mathbf{u}$$

De l'équation 14.6 pour  $\mathbf{F}^{-1}$  :

$$\mathbf{C}^{-1} = \mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} = (\mathbf{I} - \nabla \mathbf{u}) \cdot (\mathbf{I} - \nabla \mathbf{u})^\top \simeq \mathbf{I} - 2\gamma(\mathbf{u})$$

De même :

$$J = \det \mathbf{F} \simeq 1 + \nabla \cdot \mathbf{u} \quad \text{d'où } p = -k(J - 1) \simeq -k(\nabla \cdot \mathbf{u})$$

et on a ainsi, toujours en négligeant les termes d'ordre 2, le modèle néo-hookéan compressible devient :

$$\begin{aligned}
\mathbf{S} &= \mu(\mathbf{I} - \mathbf{C}^{-1}) + \frac{\lambda}{2}(\ln I_3)\mathbf{C}^{-1} \\
&\simeq \mu(\mathbf{I} - \mathbf{I} + 2\boldsymbol{\gamma}(\mathbf{u})) + \frac{\lambda}{2}2(\ln J)(\mathbf{I} - 2\boldsymbol{\gamma}(\mathbf{u})) \\
&\simeq 2\mu\boldsymbol{\gamma}(\mathbf{u}) + \lambda \ln(1 + \nabla \cdot \mathbf{u})(\mathbf{I} - 2\boldsymbol{\gamma}(\mathbf{u})) \\
&\simeq 2\mu\boldsymbol{\gamma}(\mathbf{u}) + \lambda \left( \nabla \cdot \mathbf{u} - \frac{(\nabla \cdot \mathbf{u})^2}{2} + \dots \right) (\mathbf{I} - 2\boldsymbol{\gamma}(\mathbf{u})) \\
&\simeq 2\mu\boldsymbol{\gamma}(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I}
\end{aligned}$$

De plus, pour le tenseur des contraintes, on a :

$$\begin{aligned}
\boldsymbol{\sigma} &= \frac{1}{J}\mathbf{F} \cdot \mathbf{S} \cdot \mathbf{F}^\top \\
&\simeq \left( \frac{1}{1 + \nabla \cdot \mathbf{u}} \right) (\mathbf{I} + \nabla \mathbf{u})(2\mu\boldsymbol{\gamma}(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I})(\mathbf{I} + \nabla \mathbf{u})^\top \\
&\simeq (1 - \nabla \cdot \mathbf{u} + \dots)(\mathbf{I} + \nabla \mathbf{u})(2\mu\boldsymbol{\gamma}(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I})(\mathbf{I} + \nabla \mathbf{u})^\top \\
&\simeq 2\mu\boldsymbol{\gamma}(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I}
\end{aligned}$$

en négligeant, encore une fois, les termes d'ordre 2. On montre ainsi que dans le cas des petites déformations,  $\boldsymbol{\sigma} \simeq \mathbf{S}$  qui correspond au modèle utilisé au chapitre 11. ♦

#### 14.4.1 Limite incompressible

Lors de l'étude des problèmes d'élasticité linéaire en petites déformations des chapitres 11 et 13, nous avons pu constater que la limite incompressible pose des problèmes numériques qui ont été surmontés par l'emploi de méthodes mixtes judicieusement choisies. C'est encore le cas en grandes déformations. Les matériaux incompressibles (ou quasi incompressibles) ont la particularité de se déformer tout en préservant le volume. Cela se traduit par la condition  $\det \mathbf{F} = 1$  si le matériau est parfaitement incompressible, ou plus généralement par  $\det \mathbf{F} \simeq 1$  dans le cas dit quasi-incompressible. Pour les petites déformations, cette condition est équivalente à  $\nabla \cdot \mathbf{u} \simeq 0$ , expression que nous connaissons bien. Il semble donc naturel lorsque l'on étudie de tels matériaux de découpler les déformations volumiques des autres déformations dites isochoriques. Une façon d'y arriver est d'introduire un nouveau tenseur de déformation :

$$\hat{\mathbf{F}} = (\det \mathbf{F})^{-1/3} \mathbf{F} = J^{-1/3} \mathbf{F}$$

dont le déterminant est de toute évidence 1. On obtient de cette manière une décomposition du tenseur de déformation  $\mathbf{F} = J^{1/3} \hat{\mathbf{F}}$  en une partie volumique et une partie isochorique. On obtient également une décomposition du tenseur de Cauchy-Green :

$$\mathbf{C} = J^{2/3} \hat{\mathbf{F}}^\top \hat{\mathbf{F}} = J^{2/3} \hat{\mathbf{C}} = I_3^{1/3} \hat{\mathbf{C}} \text{ ou encore } \hat{\mathbf{C}} = I_3^{-1/3} \mathbf{C}$$

On a de plus que  $\det \hat{\mathbf{C}} = 1$ . On exprimera dorénavant le potentiel  $\Psi$  en fonction des deux premiers invariants de  $\hat{\mathbf{C}}$  c.-à-d. :

$$J_1 = \text{tr}(\hat{\mathbf{C}}) = I_1 I_3^{-1/3}$$

$$J_2 = \frac{1}{2} \left( J_1^2 - \hat{\mathbf{C}} : \hat{\mathbf{C}} \right) = I_2 I_3^{-2/3}$$

mais aussi de  $J$  puisque  $J_3 = \det \hat{\mathbf{C}} = 1$ . Dans les modèles que nous verrons plus loin, le potentiel d'énergie sera plus précisément de la forme :

$$\Psi = \Psi_0(J_1, J_2) + \frac{1}{2} k (J - 1)^2$$

où  $\Psi_0(J_1, J_2)$  est un polynôme en  $J_1$  et  $J_2$  et  $k$  est le module de compressibilité. Si  $k$  est grand, on force ainsi l'incompressibilité. On a donc :

$$\mathbf{S} = 2 \frac{\partial \Psi}{\partial \mathbf{C}} = 2 \frac{\partial \Psi_0}{\partial \mathbf{C}} + 2k(J-1) \frac{\partial J}{\partial \mathbf{C}} = 2 \left( \frac{\partial \Psi_0}{\partial J_1} \frac{\partial J_1}{\partial \mathbf{C}} + \frac{\partial \Psi_0}{\partial J_2} \frac{\partial J_2}{\partial \mathbf{C}} \right) + k(J-1) J \mathbf{C}^{-1} \quad (14.22)$$

Rappelons de plus que :

$$p = -\frac{1}{3} \text{tr}(\boldsymbol{\sigma}) = -\frac{1}{3J} \text{tr}(\mathbf{F} \cdot \mathbf{S} \cdot \mathbf{F}^\top) = -\frac{1}{3J} \text{tr}(\mathbf{S} \cdot \mathbf{C})$$

en vertu des propriétés de la trace d'un produit B.5. Or, en multipliant à droite par  $\mathbf{C}$  l'expression (14.22), on constate que les scalaires  $\frac{\partial \Psi_0}{\partial J_1}$  et  $\frac{\partial \Psi_0}{\partial J_2}$  n'ont qu'un effet multiplicatif sur la trace des tenseurs  $\frac{\partial J_1}{\partial \mathbf{C}} \cdot \mathbf{C}$  et  $\frac{\partial J_2}{\partial \mathbf{C}} \cdot \mathbf{C}$  on vérifie facilement (voir les exercices de fin de chapitre) à l'aide des expressions B.20 que :

$$\text{tr} \left( \frac{\partial J_1}{\partial \mathbf{C}} \cdot \mathbf{C} \right) = \text{tr} \left( \frac{\partial J_2}{\partial \mathbf{C}} \cdot \mathbf{C} \right) = 0, \text{ de sorte que } \text{tr}(\mathbf{S} \cdot \mathbf{C}) = 3k(J-1)J$$

et par la suite :

$$p = -\frac{1}{3J} \text{tr}(\mathbf{S} \cdot \mathbf{C}) = -k(J-1)$$

Le tenseur de Piola-Kirchoff s'écrit donc :

$$\mathbf{S} = 2 \left( \frac{\partial \Psi_0}{\partial J_1} \frac{\partial J_1}{\partial \mathbf{C}} + \frac{\partial \Psi_0}{\partial J_2} \frac{\partial J_2}{\partial \mathbf{C}} \right) - p J \mathbf{C}^{-1} = \mathbf{S}' - p J \mathbf{C}^{-1} \quad (14.23)$$

Dans les développements qui suivent, on aura besoin des dérivées des invariants (calculées à l'annexe B) pour obtenir les tenseurs de Piola-Kirchoff et d'élasticité.

### Modèle de Mooney-Rivlin incompressible

Pour le modèle dit de Mooney-Rivlin<sup>3</sup>, le potentiel  $\Psi$  peut alors s'écrire :

$$\Psi = c_1(J_1 - 3) + c_2(J_2 - 3) + \frac{1}{2}k(J - 1)^2$$

de sorte que :

$$\mathbf{S} = 2 \left( c_1 \frac{\partial J_1}{\partial \mathbf{C}} + c_2 \frac{\partial J_2}{\partial \mathbf{C}} \right) - pJ\mathbf{C}^{-1}$$

Le tenseur  $\mathbf{C}$  s'écrit quant à lui :

$$\mathbf{C} = 4 \frac{\partial^2 \Psi}{\partial \mathbf{C}^2} = 4 \left( c_1 \frac{\partial^2 J_1}{\partial \mathbf{C}^2} + c_2 \frac{\partial^2 J_2}{\partial \mathbf{C}^2} - p \frac{\partial^2 J}{\partial \mathbf{C}^2} \right)$$

En développant, le tenseur de Piola-Kirchoff pour le modèle de Mooney-Rivlin s'écrit donc :

$$\begin{aligned} \mathbf{S} &= 2c_1 \left( \frac{\partial I_1}{\partial \mathbf{C}} I_3^{-1/3} - \frac{1}{3} I_1 I_3^{-4/3} \frac{\partial I_3}{\partial \mathbf{C}} \right) + 2c_2 \left( \frac{\partial I_2}{\partial \mathbf{C}} I_3^{-2/3} - \frac{2}{3} I_2 I_3^{-5/3} \frac{\partial I_3}{\partial \mathbf{C}} \right) - pJ\mathbf{C}^{-1} \\ &= 2c_1 I_3^{-1/3} \left( \mathbf{I} - \frac{1}{3} I_1 \mathbf{C}^{-1} \right) + 2c_2 I_3^{-2/3} \left( I_1 \mathbf{I} - \mathbf{C} - \frac{2}{3} I_2 \mathbf{C}^{-1} \right) - pJ\mathbf{C}^{-1} \\ &= \mathbf{S}' - pJ\mathbf{C}^{-1} \end{aligned}$$

### Modèle néo-hookéen incompressible

Le modèle néo-hookéen est un cas particulier du modèle de Mooney-Rivlin obtenu en posant  $\mu = 2c_1$  et  $c_2 = 0$  de sorte que :

$$\mathbf{S} = \mu I_3^{-1/3} \left( \mathbf{I} - \frac{1}{3} I_1 \mathbf{C}^{-1} \right) - pJ\mathbf{C}^{-1}$$

#### ◆ Remarque 14.3

Le tenseur de Cauchy peut alors s'écrire :

$$\boldsymbol{\sigma} = \frac{1}{J} \mathbf{F} \cdot \mathbf{S} \cdot \mathbf{F}^\top = \frac{1}{J} \mathbf{F} \cdot (\mathbf{S}' - pJ\mathbf{C}^{-1}) \cdot \mathbf{F}^\top = \frac{1}{J} \mathbf{F} \cdot \mathbf{S}' \cdot \mathbf{F}^\top - p\mathbf{I} = \mu J^{-5/3} \left( \mathbf{B} - \frac{I_1}{3} \mathbf{I} \right) - p\mathbf{I}$$

où  $\mathbf{B}$  est le tenseur de Finger ( $\mathbf{B} = \mathbf{F} \cdot \mathbf{F}^\top$ ). De plus, en utilisant les mêmes approximations qu'à la page 301, on retrouve le modèle élastique linéaire dans le cas des petites déformations. En effet,

3. Le britano-américain Ronald Samuel Rivlin (1915—2005) était mathématicien, physicien, rhéologue et surtout, spécialiste des caoutchoucs. Melvin Mooney (1893—1968) était un rhéologue américain, concepteur de nombreux équipements spécialement utilisés pour le caoutchouc.

on a :

$$\begin{aligned}
\frac{1}{J} \mathbf{F} \cdot \mathbf{S}' \cdot \mathbf{F}^\top &\simeq \mu J^{-5/3} (\mathbf{I} + \nabla \mathbf{u}) \cdot \left( \mathbf{I} - \frac{1}{3} (3 + 2\nabla \cdot \mathbf{u}) (\mathbf{I} - 2\gamma(\mathbf{u})) \right) \cdot (\mathbf{I} + \nabla \mathbf{u})^\top \\
&\simeq \mu (1 + \nabla \cdot \mathbf{u})^{-5/3} \cdot (\mathbf{I} + \nabla \mathbf{u}) \cdot \left( -\frac{2}{3} (\nabla \cdot \mathbf{u}) \mathbf{I} + 2\gamma(\mathbf{u}) \right) \cdot (\mathbf{I} + \nabla \mathbf{u})^\top \\
&\simeq \mu \left( 1 - \frac{5}{3} \nabla \cdot \mathbf{u} + \dots \right) \cdot \left( -\frac{2}{3} (\nabla \cdot \mathbf{u}) \mathbf{I} + 2\gamma(\mathbf{u}) \right) \\
&\simeq 2\mu \gamma(\mathbf{u}) - \frac{2\mu}{3} (\nabla \cdot \mathbf{u}) \mathbf{I}
\end{aligned}$$

qui est bien le modèle linéaire présenté à l'équation (13.4).  $\blacklozenge$

## 14.5 Formulations variationnelles

Nous avons en main tous les outils pour obtenir la formulation variationnelle ou plus précisément les formulations variationnelles. Nous considérerons en effet différentes formulations dites en déplacements ou encore mixtes. Nous écrirons également ces formulations sur les géométries déformée et non déformée.

L'équation d'équilibre sur la configuration déformée  $\Omega^t$  s'écrit :

$$-\nabla \cdot \boldsymbol{\sigma} = \mathbf{r} \quad \text{dans } \Omega^t$$

On multiplie ensuite par une fonction test  $\mathbf{w}$  et on intègre pour obtenir<sup>4</sup> :

$$\int_{\Omega^t} \boldsymbol{\sigma} : \nabla \mathbf{w} \, dv - \int_{\Gamma^t} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \mathbf{w} \, da = \int_{\Omega^t} \mathbf{r} \cdot \mathbf{w} \, dv$$

Les conditions aux limites sont également plus complexes en grandes déformations. Comme auparavant, si on impose des conditions de Dirichlet sur le déplacement  $\mathbf{u}$  sur une partie de la frontière  $\Gamma_D^t$ , alors les fonctions tests  $\mathbf{w}$  s'y annulent. On peut aussi imposer des conditions de type Neumann sur une partie  $\Gamma_N$  de la frontière de même que des conditions de contact frottant ou non sur une autre partie  $\Gamma_C$ . On parlera de contact lorsque le domaine  $\Omega^t$  rencontre un obstacle rigide ou lui-même déformable. Nous considérerons les quatre cas suivants :

1. Condition de Dirichlet en déplacement :  $\mathbf{u} = \mathbf{g}$  sur  $\Gamma_D^t$ . Les fonctions tests  $\mathbf{w}$  s'annuleront alors sur  $\Gamma_D^t$ . On peut aussi bien sûr imposer une condition aux limites sur une des composantes de  $\mathbf{u}$  seulement.

---

4. Le symbole « : » désigne le produit doublement contracté de deux tenseurs (voir l'annexe B).

2. Condition de Neumann :  $\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{h}$  sur  $\Gamma_N^t$ . Ici encore, on peut n'imposer qu'une seule des composantes.
3. Pression suiveuse :  $\boldsymbol{\sigma} \cdot \mathbf{n} = P\mathbf{n}$  sur  $\Gamma_P^t$ . Il s'agit là d'une condition fréquemment utilisée en pratique et qui nécessitera un traitement particulier. On remarque la présence de la normale à la géométrie déformée qui varie dans le temps, d'où le nom de pression suiveuse. Nous y reviendrons plus loin.
4. Condition de contact frottant ou non sur  $\Gamma_C^t$ . Un corps en déformation est susceptible d'entrer en contact avec un autre corps, lui-même déformable ou non et de s'y frotter. Les applications sont fort nombreuses mais la physique du problème est très complexe et l'implémentation de méthodes numériques demande beaucoup de soin. Mentionnons tout de même quelques mots à ce sujet. On décompose le tenseur des contraintes et la fonction test suivant leur composantes normale et tangentielle de la manière suivante :

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \sigma_n \mathbf{n} + \boldsymbol{\sigma}_t \quad \text{et} \quad \mathbf{w} = (\mathbf{w} \cdot \mathbf{n})\mathbf{n} + \mathbf{w}_t$$

La formulation variationnelle devient :

$$\int_{\Omega^t} \boldsymbol{\sigma} : \nabla \mathbf{w} \, dv = \int_{\Omega^t} \mathbf{r} \cdot \mathbf{w} \, dv + \int_{\Gamma_N^t} \mathbf{h} \cdot \mathbf{w} \, da + \int_{\Gamma_P^t} P\mathbf{n} \cdot \mathbf{w} \, da + \int_{\Gamma_C^t} (\sigma_n(\mathbf{w} \cdot \mathbf{n}) + \boldsymbol{\sigma}_t \cdot \mathbf{w}_t) \, da \quad (14.24)$$

La composante  $\sigma_n$  permettra de contrôler l'interpénétration des corps en contact tandis que  $\boldsymbol{\sigma}_t$  nous permettra d'imposer la force de frottement. Nous reviendrons plus loin sur cette formulation variationnelle. Notons que les dérivées dans cette expression sont effectuées par rapport à la configuration actuelle  $\Omega^t$  c.-à-d. par rapport aux variables  $x_i$  et que de plus, cette configuration est *a priori* inconnue. Nous allons ramener cette formulation variationnelle sur la configuration initiale  $\Omega^0$ . On distinguera les intégrales volumiques des intégrales surfaciques, même si le traitement est semblable dans les deux cas.

### 14.5.1 Intégrales volumiques

Pour se ramener sur la configuration initiale, un changement de variables classique est nécessaire. Il est quand même bon de revoir les détails.

#### ♣ Lemme 14.12

$$\int_{\Omega^t} \mathbf{r} \cdot \mathbf{w} \, dv = \int_{\Omega^0} \mathbf{r} \cdot \mathbf{w} \, J \, dV \quad \text{et} \quad \int_{\Omega^t} \boldsymbol{\sigma} : \nabla \mathbf{w} \, dv = \int_{\Omega^0} \boldsymbol{\Pi} : \nabla_X \mathbf{w} \, dV \quad (14.25)$$

#### Démonstration :

La première égalité est la formule de changement de variables classique. Pour la deuxième, on rappelle que le tenseur de contraintes  $\boldsymbol{\sigma}$  et le second tenseur de Piola-Kirchoff vérifie la relation 14.20.

De l'équation 14.5 et des propriétés de la trace et du produit doublement contracté B.5 et B.9, on a :

$$\begin{aligned}
\int_{\Omega^t} \boldsymbol{\sigma} : \nabla \mathbf{w} \, dv &= \int_{\Omega^0} \left( \frac{\mathbf{F} \cdot \mathbf{S} \cdot \mathbf{F}^\top}{J} \right) : (\nabla_X \mathbf{w} \cdot \mathbf{F}^{-1}) J \, dV \\
&= \int_{\Omega^0} \text{tr} \left[ (\mathbf{F} \cdot \mathbf{S} \cdot \mathbf{F}^\top) \cdot (\nabla_X \mathbf{w} \cdot \mathbf{F}^{-1})^\top \right] \, dV \\
&= \int_{\Omega^0} \text{tr} \left[ \mathbf{F} \cdot \mathbf{S} \cdot \mathbf{F}^\top \cdot \mathbf{F}^{-\top} \cdot \nabla_X^\top \mathbf{w} \right] \, dV \\
&= \int_{\Omega^0} (\mathbf{F} \cdot \mathbf{S}) : \nabla_X \mathbf{w} \, dV
\end{aligned}$$

On reconnaît ainsi le premier tenseur de Piola-Kirchoff. ♣

### 14.5.2 Intégrales surfaciques

Pour obtenir une formulation variationnelle complète, il nous faut transformer les intégrales surfaciques (provenant des conditions aux limites de type Neumann) de la configuration déformée à la configuration initiale. Il faudra voir aussi comment se transforme la normale. Typiquement, on devra transformer des expressions de la forme :

$$\int_{\Gamma} f \, da = \int_{\Gamma^0} f \frac{da}{dA} \, dA = \int_{\Gamma^0} f J_s \, dA$$

où le jacobien surfacique  $J_s$  est donné par la formule 14.10.

Par exemple, l'imposition des conditions aux limites naturelles nous amène à considérer :

$$\int_{\Gamma_C^t} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \mathbf{w} \, da = \int_{\Gamma_C^0} \boldsymbol{\sigma} \cdot \left( \mathbf{F}^{-\top} \cdot \mathbf{N} \frac{J}{J_s} \right) \cdot \mathbf{w} \, J_s \, dA = \int_{\Gamma_C^0} (\boldsymbol{\Pi} \cdot \mathbf{N}) \cdot \mathbf{w} \, dA$$

où on a utilisé la formule de Nanson et la relation  $\boldsymbol{\Pi} = J \boldsymbol{\sigma} \cdot \mathbf{F}^{-\top}$ .

### 14.5.3 Pression suiveuse

Les pressions dites « suiveuses » sont parmi les plus importantes conditions aux limites dans les problèmes en grandes déformations. Rappelons qu'il s'agit d'une condition de la forme  $\boldsymbol{\sigma} \cdot \mathbf{n} = P \mathbf{n}$ , d'où leur qualité de suiveuse puisque cette force surfacique suit la direction de la normale à la configuration déformée (contrairement à une condition de Neumann classique). On doit donc ajouter au second membre un terme de la forme (voir la relation 14.24) :

$$P_s = \int_{\Gamma_P^t} P \mathbf{n} \cdot \mathbf{w} \, da = \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) J \, dA$$

expression obtenue à l'aide de la formule de Nanson 14.8.

#### 14.5.4 Formulation variationnelle sur la configuration initiale

Dans un premier temps, nous négligerons les forces de contact et nous nous limiterons aux conditions de Neumann usuelles et aux pressions suiveuses. La formulation variationnelle 14.24 se réduit à :

$$\int_{\Omega^t} \boldsymbol{\sigma} : \nabla \mathbf{w} \, dv = \int_{\Gamma_N^t} \mathbf{h} \cdot \mathbf{w} \, da + \int_{\Gamma_P^t} P \mathbf{n} \cdot \mathbf{w} \, da + \int_{\Omega^t} \mathbf{r} \cdot \mathbf{w} \, dv$$

qui équivaut à résoudre sur la configuration déformée les équations :

$$\begin{cases} -\nabla \cdot \boldsymbol{\sigma} = \mathbf{r} & \text{dans } \Omega^t \\ \boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{h} & \text{sur } \Gamma_N^t \\ \boldsymbol{\sigma} \cdot \mathbf{n} = P \mathbf{n} & \text{sur } \Gamma_P^t \end{cases}$$

En se servant des résultats précédents, on peut reporter cette formulation variationnelle sur la configuration non déformée :

$$\begin{aligned} \int_{\Omega^0} (\mathbf{F} \cdot \mathbf{S}) : \nabla_X \mathbf{w} \, dV &= \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} J_s \, dA + \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) J \, dA + \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} J \, dV \\ &= \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA + \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) J \, dA + \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV \end{aligned}$$

où  $\mathbf{r}_0 = J \mathbf{r}$  et  $\mathbf{h}_0 = J_s \mathbf{h}$ . La formulation variationnelle sur la configuration non déformée est donc :

$$\int_{\Omega^0} \boldsymbol{\Pi} : \nabla_X \mathbf{w} \, dV = \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA + \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) J \, dA + \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV \quad (14.26)$$

et le problème équivaut à résoudre :

$$\begin{cases} -\nabla \cdot \boldsymbol{\Pi} = \mathbf{r}_0 & \text{dans } \Omega^0 \quad (\mathbf{r}_0(\mathbf{X}) = J \mathbf{r}(\mathbf{x})) \\ \boldsymbol{\Pi} \cdot \mathbf{N} = \mathbf{h}_0 & \text{sur } \Gamma_N^0 \quad (\mathbf{h}_0(\mathbf{X}) = J_s \mathbf{h}(x)) \\ \boldsymbol{\Pi} \cdot \mathbf{N} = P J (\mathbf{F}^{-\top} \cdot \mathbf{N}) & \text{sur } \Gamma_P^0 \end{cases}$$

sur la configuration initiale (non déformée). C'est donc le premier tenseur de Piola-Kirchoff qui apparaît dans la configuration initiale et qui joue le même rôle que le tenseur  $\boldsymbol{\sigma}$  sur la configuration déformée. Les autres types de conditions aux limites seront abordés un peu plus loin. Une autre forme équivalente est la suivante :

$$\int_{\Omega^0} \mathbf{S} : (\mathbf{F}^\top \cdot \nabla_X \mathbf{w}) \, dV = \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA + \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) J \, dA + \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV \quad (14.27)$$

### Formulation variationnelle mixte pour matériaux quasi-incompressibles

Pour simplifier l'exposé, nous supposons qu'il n'y a aucune pression suiveuse imposée. Si on utilise la décomposition 14.23, la formulation variationnelle 14.26 devient :

$$\int_{\Omega^0} (\mathbf{F} \cdot (\mathbf{S}' - pJ\mathbf{C}^{-1})) : \nabla_X \mathbf{w} \, dV = \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA + \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV$$

qui sous sa forme actuelle, nous amène directement à une formulation mixte puisqu'on doit jumeler la dernière équation à la condition :

$$p = -k(J - 1) \quad \text{ou encore} \quad \frac{p}{k} + (J - 1) = 0$$

qui devient sous forme variationnelle :

$$\int_{\Omega^0} \left( (J - 1) + \frac{1}{k} p \right) q \, dV = 0 \quad \forall q$$

On remarque de plus que  $\mathbf{F} \cdot \mathbf{C}^{-1} = \mathbf{F}^{-\top}$  et une formulation mixte en grandes déformations aura donc la forme du système suivant :

$$\begin{aligned} \int_{\Omega^0} (\mathbf{F} \cdot \mathbf{S}') : \nabla_X \mathbf{w} \, dV - \int_{\Omega^0} pJ\mathbf{F}^{-\top} : \nabla_X \mathbf{w} \, dV &= \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA + \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV \\ \int_{\Omega^0} (J - 1)q \, dV + \int_{\Omega^0} \frac{1}{k} pq \, dV &= 0 \end{aligned}$$

ou en utilisant encore les propriétés du produit doublement contracté :

$$\left\{ \begin{aligned} \int_{\Omega^0} \mathbf{S}' : (\mathbf{F}^\top \cdot \nabla_X \mathbf{w}) \, dV - \int_{\Omega^0} pJ\mathbf{F}^{-\top} : \nabla_X \mathbf{w} \, dV &= \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA + \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV \\ \int_{\Omega^0} (J - 1)q \, dV + \int_{\Omega^0} \frac{1}{k} pq \, dV &= 0 \end{aligned} \right. \quad (14.28)$$

Cette formulation est non linéaire et devra donc éventuellement être linéarisée.

## 14.6 Linéarisation et méthode de Newton

La formulation variationnelle du problème étant non linéaire, la méthode de Newton semble la plus appropriée pour sa résolution et elle est fréquemment utilisée. On doit donc effectuer une linéarisation.

### 14.6.1 Formulation en déplacement seulement

Cette formulation, exprimée par la relation 14.26, est pertinente dans le cas où le matériau est loin de la limite incompressible. Dans le cas incompressible (ou quasi-incompressible), on utilisera avantageusement la formulation mixte de la section 14.6.2. On pose donc :

$$R(\mathbf{u}, \mathbf{w}) = \int_{\Omega^0} \mathbf{F}(\mathbf{u}) \cdot \mathbf{S}(\mathbf{u}) : \nabla_X \mathbf{w} \, dV - \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA - \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) \, dA - \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV$$

et partant de  $\mathbf{u}_0$ , on cherche une correction  $\delta_u$  de sorte que  $R(\mathbf{u}_0 + \delta_u, \mathbf{w}) = 0$ . On constate aisément qu'il y a deux termes à linéariser soit le terme incluant la contribution du tenseur de Piola-Kirchoff et le terme de pression suiveuse.

#### Linéarisation du matériau

Le premier terme à linéariser fait intervenir le tenseur de Piola-Kirchoff :

$$\int_{\Omega^0} \mathbf{F}(\mathbf{u}) \cdot \mathbf{S}(\mathbf{u}) : \nabla_X \mathbf{w} \, dV = \int_{\Omega^0} \mathbf{S}(\mathbf{u}) : (\mathbf{F}^\top(\mathbf{u}) \cdot \nabla_X \mathbf{w}) \, dV$$

en vertu de B.10. La dérivée est donc :

$$\int_{\Omega^0} \left( \frac{\partial \mathbf{S}}{\partial \mathbf{u}} \cdot \delta_u \right) : (\mathbf{F}^\top \cdot \nabla_X \mathbf{w}) \, dV + \int_{\Omega^0} \mathbf{S} : \left( \left( \frac{\partial \mathbf{F}^\top}{\partial \mathbf{u}} \cdot \delta_u \right) \cdot \nabla_X \mathbf{w} \right) \, dV$$

#### ♣ Lemme 14.13

$$\frac{\partial \mathbf{S}}{\partial \mathbf{u}} \cdot \delta_u = \frac{\partial \mathbf{S}}{\partial \mathbf{E}} : \left( \frac{\partial \mathbf{E}}{\partial \mathbf{u}} \cdot \delta_u \right) = \mathcal{C} : \frac{(\mathbf{F}^\top \cdot \nabla_X \delta_u + \nabla_X^\top \delta_u \cdot \mathbf{F})}{2} = \mathcal{C} : (\mathbf{F}^\top \cdot \nabla_X \delta_u) \quad (14.29)$$

$$\frac{\partial \mathbf{F}^\top}{\partial \mathbf{u}} \cdot \delta_u = \nabla_X^\top(\delta_u)$$

**Démonstration** : Le résultat suit de la définition de  $\mathcal{C}$  et de la relation B.23. ♣

La linéarisation du premier terme devient donc :

$$\int_{\Omega^0} \mathbf{S}(\mathbf{u}_0) : (\nabla_X^\top(\delta_u) \cdot \nabla_X \mathbf{w}) \, dV + \int_{\Omega^0} \left( \mathcal{C}(\mathbf{u}_0) : (\mathbf{F}^\top(\mathbf{u}_0) \cdot \nabla_X(\delta_u)) \right) : (\mathbf{F}^\top(\mathbf{u}_0) \cdot \nabla_X \mathbf{w}) \, dV$$

### Linéarisation de la pression suiveuse

Rappelons qu'une pression suiveuse fait intervenir un terme de la forme :

$$P_s = \int_{\Gamma_P^i} P \mathbf{n} \cdot \mathbf{w} \, da = \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) J \, dA$$

qu'il faut aussi linéariser par rapport au déplacement  $\mathbf{u}$ . On constate que seuls  $J$  et  $\mathbf{F}^{-\top}$  dépendent de  $\mathbf{u}$ . Il faudra donc évaluer leur variation par rapport à  $\mathbf{u}$  c.-à-d. :

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{u}} \cdot \delta \mathbf{u} &= \frac{\partial J}{\partial \mathbf{F}} : \left( \frac{\partial \mathbf{F}}{\partial \mathbf{u}} \cdot \delta \mathbf{u} \right) \\ \frac{\partial \mathbf{F}^{-\top}}{\partial \mathbf{u}} \cdot \delta \mathbf{u} &= \frac{\partial \mathbf{F}^{-\top}}{\partial \mathbf{F}} : \left( \frac{\partial \mathbf{F}}{\partial \mathbf{u}} \delta \mathbf{u} \right) \end{aligned}$$

Rappelons qu'en vertu des relations B.21, B.15 et B.17, on a immédiatement que :

$$\frac{\partial \mathbf{F}}{\partial \mathbf{u}} \cdot \delta \mathbf{u} = \nabla_X \delta \mathbf{u}, \quad \frac{\partial J}{\partial \mathbf{F}} : \delta \mathbf{F} = J \mathbf{F}^{-\top} : \delta \mathbf{F} \quad \text{et} \quad \frac{\partial \mathbf{F}^{-\top}}{\partial \mathbf{F}} : \delta \mathbf{F} = -\mathbf{F}^{-\top} \cdot (\delta \mathbf{F})^\top \cdot \mathbf{F}^{-\top} = \mathcal{K} : \delta \mathbf{F} \quad (14.30)$$

de sorte que :

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{u}} \cdot \delta \mathbf{u} &= J \mathbf{F}^{-\top} : \nabla_X \delta \mathbf{u} \\ \frac{\partial \mathbf{F}^{-\top}}{\partial \mathbf{u}} \cdot \delta \mathbf{u} &= -\mathbf{F}^{-\top} \cdot (\nabla_X \delta \mathbf{u})^\top \cdot \mathbf{F}^{-\top} = \mathcal{K} : \nabla_X \delta \mathbf{u} \end{aligned}$$

En dérivant le terme de pression suiveuse, on doit ajouter à la formulation variationnelle (la matrice tangente) une contribution de la forme :

$$\begin{aligned} \frac{\partial P_s}{\partial \mathbf{u}} \cdot \delta \mathbf{u} &= \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ \left( \frac{\partial J}{\partial \mathbf{u}} \delta \mathbf{u} \right) (\mathbf{F}^{-\top} \cdot \mathbf{N}) + J \left( \frac{\partial \mathbf{F}^{-\top}}{\partial \mathbf{u}} \delta \mathbf{u} \right) \cdot \mathbf{N} \right] dA \\ &= \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ \left( \mathbf{F}^{-\top} : \nabla_X \delta \mathbf{u} \right) (\mathbf{F}^{-\top} \cdot \mathbf{N}) - \left( \mathbf{F}^{-\top} \cdot (\nabla_X \delta \mathbf{u})^\top \cdot \mathbf{F}^{-\top} \right) \cdot \mathbf{N} \right] J \, dA \\ &= \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ \left( \mathbf{F}^{-\top} : \nabla_X \delta \mathbf{u} \right) (\mathbf{F}^{-\top} \cdot \mathbf{N}) + (\mathcal{K} : \nabla_X \delta \mathbf{u}) \cdot \mathbf{N} \right] J \, dA \\ &= \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ \left( \mathbf{F}^{-\top} \otimes \mathbf{F}^{-\top} + \mathcal{K} \right) : \nabla_X \delta \mathbf{u} \right] \cdot \mathbf{N} J \, dA \quad (\text{en vertu de la relation B.8}) \end{aligned}$$

qui n'est pas symétrique. La perte de symétrie due à la linéarisation des pressions suiveuses exige de mettre en mémoire une matrice complète non symétrique, au lieu d'une demi matrice seulement.

Cette perte de symétrie a aussi des conséquences sur le choix des méthodes itératives de résolution des systèmes linéaires qui en résultent puisque certaines méthodes comme le gradient conjugué ne fonctionnent que pour des matrices symétriques.

Pour contourner partiellement cette difficulté, on peut symétriser la contribution de la pression suivieuse à la matrice tangente. Si on dénote  $M_P$  cette matrice, on peut en effet la remplacer par  $(M_P + M_P^\top)/2$ . Le prix à payer est que dans ces conditions, nous ne sommes plus en présence d'une méthode de Newton et que la convergence n'est plus quadratique. En fait l'expérience montre qu'il faut alors charger la pression suivieuse par incréments progressifs, ce qui augmente sensiblement le coût de calcul. Ce qui a été gagné en espace mémoire est perdu en temps de calcul..

► **Exemple 14.1** On considère dans cet exemple une tige de section carrée faite d'un matériau hyperélastique. On fixe l'extrémité gauche de la tige de sorte que les déplacements y soient nuls c.-à-d.  $\mathbf{u} = 0$ . Sur la paroi supérieure de cette tige, on impose d'abord une condition de type Neumann  $\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{h}$  qui comme on l'a vu, est équivalente à imposer  $\boldsymbol{\Pi} \cdot \mathbf{n} = J_s \mathbf{h}$  sur la configuration non déformée. ◀

### Système à résoudre

En résumé, la méthode de Newton pour une formulation en déplacement seulement requiert la résolution de systèmes linéaires de la forme :

$$\frac{\partial R(\mathbf{u}_0, \mathbf{w})}{\partial \mathbf{u}} \cdot \boldsymbol{\delta}_u = -R(\mathbf{u}_0, \mathbf{w})$$

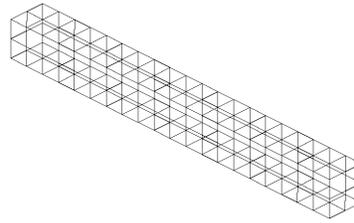
où :

$$\begin{aligned} \frac{\partial R(\mathbf{u}_0, \mathbf{w})}{\partial \mathbf{u}} \cdot \boldsymbol{\delta}_u &= \int_{\Omega^0} \mathbf{S}(\mathbf{u}_0) : \left( \nabla_X^\top(\boldsymbol{\delta}_u) \cdot \nabla_X \mathbf{w} \right) dV \\ &+ \int_{\Omega^0} \left( \mathbf{C}(\mathbf{u}_0) : \left( \mathbf{F}^\top(\mathbf{u}_0) \cdot \nabla_X(\boldsymbol{\delta}_u) \right) \right) : \left( \mathbf{F}^\top(\mathbf{u}_0) \cdot \nabla_X \mathbf{w} \right) dV \\ &- \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ \left( \mathbf{F}^{-\top} : \nabla_X \boldsymbol{\delta}_u \right) (\mathbf{F}^{-\top} \cdot \mathbf{N}) - \left( \mathbf{F}^{-\top} \cdot (\nabla_X \boldsymbol{\delta}_u)^\top \cdot \mathbf{F}^{-\top} \right) \cdot \mathbf{N} \right] J dA \end{aligned} \quad (14.31)$$

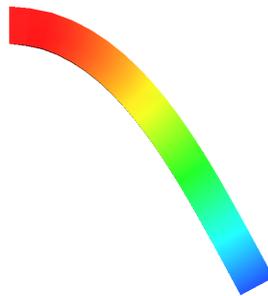
### 14.6.2 Formulation mixte en déplacements-pression

On fera l'hypothèse que le tenseur de Piola-Kirchoff ne dépend que du déplacement et de la pression sous la forme :

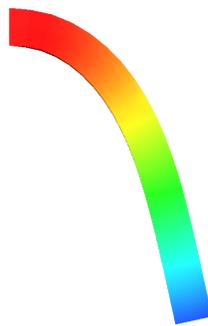
$$\mathbf{S}(\mathbf{u}, p) = \mathbf{S}(\mathbf{E}(\mathbf{u}), p)$$



a) Maillage



b) Condition de Neumann



c) Pression suivieuse

Figure 14.2 – Comparaison d’une condition de Neumann et d’une pression suivieuse

On pose ensuite :

$$\begin{aligned}
R_1((\mathbf{u}, p), \mathbf{w}) &= \int_{\Omega^0} \mathbf{S} : (\mathbf{F}^\top \cdot \nabla_X \mathbf{w}) \, dV - \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA - \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV \\
&= \int_{\Omega^0} \mathbf{S}' : (\mathbf{F}^\top \cdot \nabla_X \mathbf{w}) \, dV - \int_{\Omega^0} p J \mathbf{F}^{-\top} : \nabla_X \mathbf{w} \, dV - \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA - \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dV \\
R_2((\mathbf{u}, p), q) &= \int_{\Omega^0} \left( -(J-1) - \frac{1}{k} p \right) q \, dV
\end{aligned}$$

Partant d'une approximation  $(\mathbf{u}_0, p_0)$  du déplacement, on cherche une correction  $(\delta_u, \delta_p)$  de sorte que

$$\begin{aligned}
R_1((\mathbf{u}_0 + \delta_u, p + \delta_p), \mathbf{w}) &= 0 \\
R_2((\mathbf{u}_0 + \delta_u, p + \delta_p), q) &= 0
\end{aligned}$$

On doit donc linéariser ce système par rapport à chacune des variables  $\mathbf{u}$  et  $p$ . Pour ce faire, on a besoin de quelques lemmes.

♣ **Lemme 14.14**

$$\frac{\partial \mathbf{S}}{\partial p} q = -q J \mathbf{C}^{-1} \tag{14.32}$$

**Démonstration :**

La démonstration découle de la définition de  $\mathbf{S}$ . ♣

♣ **Lemme 14.15**

On a :

$$\begin{aligned}
\frac{\partial R_1((\mathbf{u}_0, p_0), \mathbf{w})}{\partial \mathbf{u}} \cdot \delta_u &= \int_{\Omega^0} \mathbf{S}(\mathbf{u}_0, p_0) : \left( \nabla_X^\top(\delta_u) \cdot \nabla_X \mathbf{w} \right) \, dV + \\
&\quad \int_{\Omega^0} \left( \mathbf{C}(\mathbf{u}_0, p_0) : \left( \mathbf{F}^\top(\mathbf{u}_0) \cdot \nabla_X(\delta_u) \right) \right) : \left( \mathbf{F}^\top(\mathbf{u}_0) \cdot \nabla_X \mathbf{w} \right) \, dV \\
\frac{\partial R_1((\mathbf{u}_0, p_0), \mathbf{w})}{\partial p} \delta_p &= - \int_{\Omega^0} J \delta_p \left( \mathbf{F}^{-\top}(\mathbf{u}_0) : \nabla_X \mathbf{w} \right) \, dV \\
\frac{\partial R_2((\mathbf{u}_0, p_0), q)}{\partial \mathbf{u}} \cdot \delta_u &= - \int_{\Omega^0} J q \left( \mathbf{F}^{-\top}(\mathbf{u}_0) : \nabla_X \delta_u \right) \, dV \\
\frac{\partial R_2((\mathbf{u}_0, p_0), q)}{\partial p} \delta_p &= - \int_{\Omega^0} \frac{1}{k} \delta_p q \, dV
\end{aligned}$$

où  $\mathbf{C}$  est le tenseur du quatrième ordre défini par :

$$\mathbf{C} = \frac{\partial \mathbf{S}}{\partial \mathbf{E}} = 2 \frac{\partial \mathbf{S}}{\partial \mathbf{C}}$$

**Démonstration :**

La démonstration découle des lemmes précédents. <sup>5</sup> ♣

La formulation variationnelle linéarisée est donc :

$$\left\{ \begin{array}{l} \frac{\partial R_1((\mathbf{u}_0, p_0), \mathbf{w})}{\partial \mathbf{u}} \cdot \delta_{\mathbf{u}} - \int_{\Omega^0} \delta_p J \mathbf{F}^{-\top}(\mathbf{u}_0) : \nabla_X \mathbf{w} \, dV = -R_1((\mathbf{u}_0, p_0), \mathbf{w}) \\ - \int_{\Omega^0} q J \mathbf{F}^{-\top}(\mathbf{u}_0) : \nabla_X \delta_{\mathbf{u}} \, dV - \int_{\Omega^0} \frac{1}{k} \delta_p q \, dV = -R_2((\mathbf{u}_0, p_0), q) \end{array} \right. \quad (14.33)$$

En vertu du lemme précédent, on remarque que pour les matériaux hyperélastiques, le système linéarisé est symétrique, sauf en présence d'une pression suiveuse (voir la section 14.6.1).

### 14.6.3 Formulation pénalisée

La pénalisation est une technique couramment utilisée dans le but de diminuer la taille des systèmes non linéaires à résoudre. Elle permet, dans certaines conditions, d'éliminer la variable  $p$  du système. Cette méthode est particulièrement efficace dans le cas de la discrétisation de  $p$  par des polynômes discontinus d'un élément à l'autre. Les cas les plus fréquents sont les approximations constante et linéaire par élément. On se référera au chapitre 12 où ce type de discrétisations a déjà été discuté.

Sur un élément  $K$ , la formulation variationnelle linéarisée 14.33 devient, sous forme matricielle :

$$\begin{bmatrix} A^K & B^{K\top} \\ B^K & M^K \end{bmatrix} \begin{bmatrix} \delta_{\mathbf{u}}^K \\ \delta_p^K \end{bmatrix} = - \begin{bmatrix} R_1^K \\ R_2^K \end{bmatrix}$$

où les différentes matrices élémentaires sont directement définies par 14.33. Lorsque la variable  $p$  est discontinue par élément, on peut résoudre localement pour  $p^K$ . En effet, on a :

$$\delta_p^K = -(M^K)^{-1} (R_2^K + B^K \delta_{\mathbf{u}}^K)$$

d'où l'on tire :

$$\left( A^K - B^{K\top} (M^K)^{-1} B^K \right) \delta_{\mathbf{u}}^K = -R_1^K + B^{K\top} (M^K)^{-1} R_2^K$$

---

5. L'expression  $\frac{\partial \mathbf{S}}{\partial \mathbf{E}}$  désigne un tenseur d'ordre 4 que nous noterons  $\mathbf{C}$  et dont les composantes sont :

$$c_{ijkl} = \frac{1}{2} \left( \frac{\partial S_{ij}}{\partial E_{kl}} + \frac{\partial S_{ij}}{\partial E_{lk}} \right)$$

de manière à avoir les bonnes propriétés de symétrie (voir la référence [5]).

Cette matrice et le terme de droite sont calculés directement au niveau de l'élément. Cette dernière expression n'est rien d'autre que l'algorithme d'Uzawa [20]. Pour obtenir une méthode de pénalisation classique, il suffit de poser  $p_k = 0$ , ce qui élimine la mise à jour de la pression. Le cas le plus fréquent est celui de la pression constante par élément. La matrice  $M^K$  est alors de dimension 1 sur 1 et vaut tout simplement  $\text{vol}(K)/k$ . On a ainsi :

$$\delta_p^K = \frac{-k}{\text{vol}(K)} \left( \int_K J \left( \mathbf{F}^{-\top}(\mathbf{u}_0) : \nabla_X \delta_u \right) dV + \int_K (J - 1) dV - \frac{p^K}{k} \text{vol}(K) \right)$$

## 14.7 Liens avec la géométrie différentielle intrinsèque

Les outils développés jusqu'à maintenant nous permettent d'établir quelques liens avec la géométrie différentielle intrinsèque au sens de Delfour et Zolézio [12]. On pourra notamment développer des expressions pour les variations de la pression suivieuse, de la normale  $\mathbf{n}$  et du jacobien surfacique  $J_s$  en fonction du déplacement.

### 14.7.1 Variation du jacobien

On a un lien direct entre la variation du jacobien et la divergence. En effet :

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{u}} \cdot \delta_u &= J \mathbf{F}^{-\top} : \nabla_X \delta_u \\ &= J \text{tr}(\nabla_X \delta_u \cdot \mathbf{F}^{-1}) \\ &= J \text{tr}(\nabla \delta_u) \\ &= J \nabla \cdot \delta_u \end{aligned}$$

### 14.7.2 Variation du jacobien surfacique

On sait que :

$$J_s = J \sqrt{(\mathbf{C}^{-1} \cdot \mathbf{N}) \cdot \mathbf{N}}$$

et on cherche la variation de  $J_s$  par rapport au déplacement  $\mathbf{u}$ . On a donc :

$$\frac{\partial J_s}{\partial \mathbf{u}} \cdot \delta_u = \frac{\partial J_s}{\partial \mathbf{F}} : \frac{\partial \mathbf{F}}{\partial \mathbf{u}} \delta_u = \frac{\partial J_s}{\partial \mathbf{F}} : \nabla_X \delta_u$$

On a d'une part en vertu de B.15 :

$$\frac{\partial J}{\partial \mathbf{F}} : \delta_F = J \mathbf{F}^{-\top} : \delta_F$$

et d'autre part, en utilisant B.16 et B.17 :

$$\begin{aligned}
\frac{\partial \mathbf{C}^{-1}}{\partial \mathbf{F}} : \boldsymbol{\delta}_F &= \frac{\partial (\mathbf{F}^{-1} \cdot \mathbf{F}^{-\top})}{\partial \mathbf{F}} : \boldsymbol{\delta}_F \\
&= \left( \frac{\partial \mathbf{F}^{-1}}{\partial \mathbf{F}} : \boldsymbol{\delta}_F \right) \cdot \mathbf{F}^{-\top} + \mathbf{F}^{-1} \cdot \left( \frac{\partial \mathbf{F}^{-\top}}{\partial \mathbf{F}} : \boldsymbol{\delta}_F \right) \\
&= -\mathbf{F}^{-1} \cdot \boldsymbol{\delta}_F \cdot \mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} - \mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \\
&= - \left( \left[ \mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \right]^\top + \mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \right)
\end{aligned}$$

et donc en posant  $\alpha = \sqrt{(\mathbf{C}^{-1} \cdot \mathbf{N}) \cdot \mathbf{N}}$  de sorte que  $J_s = J\alpha$ , on trouve :

$$\begin{aligned}
\frac{\partial J_s}{\partial \mathbf{F}} : \boldsymbol{\delta}_F &= \left( \frac{\partial J}{\partial \mathbf{F}} : \boldsymbol{\delta}_F \right) \alpha + \frac{J}{2\alpha} \left( \left( \frac{\partial \mathbf{C}^{-1}}{\partial \mathbf{F}} : \boldsymbol{\delta}_F \right) \cdot \mathbf{N} \right) \cdot \mathbf{N} \\
&= J \alpha \mathbf{F}^{-\top} : \boldsymbol{\delta}_F - \frac{J}{2\alpha} \left( \left( \left[ \mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \right]^\top + \mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \right) \cdot \mathbf{N} \right) \cdot \mathbf{N} \\
&= J_s \mathbf{F}^{-\top} : \boldsymbol{\delta}_F - \frac{J}{\alpha} \left( \mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \cdot \mathbf{N} \right) \cdot \mathbf{N}
\end{aligned}$$

Poursuivons plus avant, en rappelant que  $(\mathbf{A} \cdot \mathbf{x}) \cdot (\mathbf{y}) = \mathbf{x} \cdot (\mathbf{A}^\top \cdot \mathbf{y})$  :

$$\begin{aligned}
\frac{\partial J_s}{\partial \mathbf{F}} : \boldsymbol{\delta}_F &= J_s \mathbf{F}^{-\top} : \boldsymbol{\delta}_F - \frac{J}{\alpha} \left( \left( \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \cdot \mathbf{N} \right) \cdot \left( \mathbf{F}^{-\top} \cdot \mathbf{N} \right) \right) \\
&= J_s \mathbf{F}^{-\top} : \boldsymbol{\delta}_F - J\alpha \left( \left( \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \frac{\mathbf{F}^{-\top} \cdot \mathbf{N}}{\alpha} \right) \cdot \left( \frac{\mathbf{F}^{-\top} \cdot \mathbf{N}}{\alpha} \right) \right) \\
&= J_s \mathbf{F}^{-\top} : \boldsymbol{\delta}_F - J_s \left( \left( \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right) \quad \text{avec } \mathbf{n} = \frac{\mathbf{F}^{-\top} \cdot \mathbf{N}}{\|\mathbf{F}^{-\top} \cdot \mathbf{N}\|}
\end{aligned}$$

de sorte qu'en se servant de la relation 14.4, on trouve :

$$\frac{\partial J_s}{\partial \mathbf{u}} \cdot \boldsymbol{\delta}_u = J_s \left( \mathbf{F}^{-\top} : \nabla_X \boldsymbol{\delta}_u - \left( \mathbf{F}^{-\top} \cdot (\nabla_X \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right) \quad (14.34)$$

On peut encore simplifier cette expression car en effet :

$$\begin{aligned}
\frac{\partial J_s}{\partial \mathbf{u}} \cdot \boldsymbol{\delta}_u &= J_s \left( \text{tr} [\nabla_X \boldsymbol{\delta}_u \cdot \mathbf{F}^{-1}] - \left( \mathbf{F}^{-\top} \cdot (\nabla_X \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right) \\
&= J_s \left( \text{tr} [\nabla \boldsymbol{\delta}_u] - \left( (\nabla \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right) \\
&= J_s \left( \nabla \cdot \boldsymbol{\delta}_u - \left( (\nabla \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right) \\
&= J_s \nabla_\Gamma \cdot \boldsymbol{\delta}_u
\end{aligned}$$

On a ainsi retrouvé un résultat semblable à celui de la relation (B.24) mais pour le jacobien surfacique. On a du même coup introduit la *divergence tangentielle*  $\nabla_\Gamma \cdot \boldsymbol{\delta}_u$  au sens de Delfour et Zolézio [12].

#### Définition 14.5

La divergence tangentielle d'un vecteur  $\mathbf{u}$  est définie par :

$$\nabla_\Gamma \cdot \mathbf{u} = \nabla \cdot \mathbf{u} - \left( (\nabla \mathbf{u})^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \quad (14.35)$$

#### ◆ Remarque 14.4

Un cas particulier intéressant du résultat précédent est que si  $\mathbf{u} = \mathbf{n}$ , le vecteur normal à la surface, alors  $\nabla_\Gamma \cdot \mathbf{n} = \nabla \cdot \mathbf{n}$ . Il suffit en effet de vérifier que  $(\nabla \mathbf{n})^\top \cdot \mathbf{n} = 0$  en dérivant la relation  $\mathbf{n} \cdot \mathbf{n} = 1$ . ◆

### 14.7.3 Variation de la normale

De manière similaire, on peut retrouver la formule pour la variation de la normale. Rappelons que (voir 14.11) :

$$\mathbf{n} = \frac{\mathbf{F}^{-\top} \cdot \mathbf{N}}{\sqrt{(\mathbf{C}^{-1} \cdot \mathbf{N}) \cdot \mathbf{N}}} = \frac{\mathbf{F}^{-\top} \cdot \mathbf{N}}{\alpha}$$

On a ainsi :

$$\frac{\partial \mathbf{n}}{\partial \mathbf{u}} \boldsymbol{\delta}_u = \frac{\partial \mathbf{n}}{\partial \mathbf{F}} : \nabla_X \boldsymbol{\delta}_u$$

et

$$\begin{aligned}
\frac{\partial \mathbf{n}}{\partial \mathbf{F}} : \boldsymbol{\delta}_F &= \frac{\left( \frac{\partial \mathbf{F}^{-\top}}{\partial \mathbf{F}} : \boldsymbol{\delta}_F \right) \cdot \mathbf{N} \alpha - \frac{\mathbf{F}^{-\top} \cdot \mathbf{N}}{2\alpha} \left( \left( \frac{\partial \mathbf{C}^{-1}}{\partial \mathbf{F}} : \boldsymbol{\delta}_F \right) \cdot \mathbf{N} \right) \cdot \mathbf{N}}{\alpha^2} \\
&= \frac{-\alpha \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \cdot \mathbf{N} + \frac{\mathbf{F}^{-\top} \cdot \mathbf{N}}{\alpha} \left( (\mathbf{F}^{-1} \cdot \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \cdot \mathbf{N}) \cdot \mathbf{N} \right)}{\alpha^2} \\
&= \frac{-\mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \cdot \mathbf{N}}{\alpha} + \frac{\frac{\mathbf{F}^{-\top} \cdot \mathbf{N}}{\alpha} \left( (\mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{F}^{-\top} \cdot \mathbf{N}) \cdot (\mathbf{F}^{-\top} \cdot \mathbf{N}) \right)}{\alpha^2} \\
&= -\mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{n} + \mathbf{n} \left( \left( \mathbf{F}^{-\top} \cdot (\boldsymbol{\delta}_F)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right)
\end{aligned}$$

de sorte que :

$$\begin{aligned}
\frac{\partial \mathbf{n}}{\partial \mathbf{u}} \boldsymbol{\delta}_u &= -\mathbf{F}^{-\top} \cdot (\nabla_X \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} + \mathbf{n} \left( \left( \mathbf{F}^{-\top} \cdot (\nabla_X \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right) \\
&= -(\nabla \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} + \left( \left( (\nabla \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right) \mathbf{n}
\end{aligned} \tag{14.36}$$

où l'on reconnaît la dérivée matérielle de  $\mathbf{n}$  au sens de Delfour et Zolézio [12] ( $\dot{\mathbf{n}}$  dans leur notation).

#### 14.7.4 Pression suiveuse

Revenons sur la linéarisation du terme de pression suiveuse qui peut aussi s'écrire différemment. Le terme de pression suiveuse est de la forme :

$$P_s = \int_{\Gamma_P^t} P \mathbf{w} \cdot \mathbf{n} \, da = \int_{\Gamma_P^0} P \mathbf{w} \cdot \mathbf{n} J_s \, dA$$

et la linéarisation de ce terme exigera le calcul de :

$$\begin{aligned}
\frac{\partial P_s}{\partial \mathbf{u}} \cdot \boldsymbol{\delta}_u &= \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ \left( \frac{\partial \mathbf{n}}{\partial \mathbf{u}} \boldsymbol{\delta}_u \right) J_s + \mathbf{n} \left( \frac{\partial J_s}{\partial \mathbf{u}} \boldsymbol{\delta}_u \right) \right] dA \\
&= \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ \left( -(\nabla \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} + \left( \left( (\nabla \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} \right) \cdot \mathbf{n} \right) \mathbf{n} \right) + \mathbf{n} (\nabla_\Gamma \cdot \boldsymbol{\delta}_u) \right] J_s \, dA \\
&= \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ (\nabla \cdot \boldsymbol{\delta}_u) \mathbf{n} - (\nabla \boldsymbol{\delta}_u)^\top \cdot \mathbf{n} \right] J_s \, dA \\
&= \int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ (\nabla \cdot \boldsymbol{\delta}_u \mathbf{I} - (\nabla \boldsymbol{\delta}_u)^\top) \cdot \mathbf{n} \right] J_s \, dA
\end{aligned}$$

où  $J_s$  est le jacobien surfacique 14.10. On a ainsi fait apparaître le déviateur du tenseur de déformation  $\nabla \delta_u$ , ce qui semble raisonnable puisqu'un déplacement virtuel dans la direction normale ne provoque aucune variation de la normale.

## 14.8 Formulation Lagrangienne actualisée

La formulation Lagrangienne totale considérée jusqu'à maintenant peut s'avérer inadéquate pour les problèmes où la géométrie initiale  $\Omega_0$  subit de très fortes déformations. Le maillage de la géométrie initiale devient alors inapproprié pour tenir compte de ces déformations et il faut alors réactualiser la géométrie et relancer les calculs à partir cette nouvelle géométrie. Il faut bien sûr s'assurer de conserver l'historique des déformations. C'est ce que nous allons maintenant décrire.

Nous supposons donc qu'il n'est pas possible, ou à tout le moins difficile, de passer directement de la configuration  $\Omega^0$  à une nouvelle configuration notée  $\Omega^2$  en raison de déformations trop importantes. Il nous faudra donc passer par une configuration intermédiaire que nous noterons  $\Omega^1$ . Nous noterons  $\mathbf{X}_i, i = 0, 1, 2$ , les coordonnées dans la configuration  $\Omega^i$  et  $\phi_{i \rightarrow j}$  l'application qui envoie  $\Omega^i$  sur  $\Omega^j$ . Le déplacement entre les différentes géométries sera noté  $\mathbf{u}_{i \rightarrow j}$  c.-à-d.

$$\phi_{i \rightarrow j}(\mathbf{X}_i) = \mathbf{X}_j = \mathbf{X}_i + \mathbf{u}_{i \rightarrow j}$$

Ainsi, on aura par exemple  $\mathbf{X}_2 = \mathbf{X}_0 + \mathbf{u}_{0 \rightarrow 2}$  que l'on notait  $\mathbf{x} = \mathbf{X} + \mathbf{u}$  en formulation Lagrangienne totale. On peut maintenant aussi écrire :

$$\mathbf{X}_2 = \mathbf{X}_1 + \mathbf{u}_{1 \rightarrow 2} = (\mathbf{X}_0 + \mathbf{u}_{0 \rightarrow 1}) + \mathbf{u}_{1 \rightarrow 2} = \mathbf{X}_0 + \mathbf{u}_{0 \rightarrow 2}$$

et donc  $\mathbf{u}_{0 \rightarrow 2} = \mathbf{u}_{0 \rightarrow 1} + \mathbf{u}_{1 \rightarrow 2}$  ainsi que :

$$\mathbf{X}_2 = \phi_{1 \rightarrow 2}(\mathbf{X}_1) = \phi_{1 \rightarrow 2}(\phi_{0 \rightarrow 1}(\mathbf{X}_0)) = \phi_{1 \rightarrow 2} \circ \phi_{0 \rightarrow 1}(\mathbf{X}_0)$$

En notant maintenant le tenseur de déformation  $\mathbf{F}_{i \rightarrow j}$  défini par :

$$(\mathbf{F}_{i \rightarrow j})_{mn} = \frac{\partial (\mathbf{X}_j)_m}{\partial (\mathbf{X}_i)_n}$$

où  $(\mathbf{X}_j)_m$  est la m-ième composante de  $\mathbf{X}_j$ , on obtient immédiatement par composition de fonctions que :

$$\mathbf{F}_{0 \rightarrow 2} = \mathbf{F}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1} \quad \text{et donc} \quad J_{0 \rightarrow 2} = J_{1 \rightarrow 2} J_{0 \rightarrow 1}$$

### 14.8.1 Formulation en déplacement seulement

En suivant cette notation, pour passer de la configuration  $\Omega^0$  à  $\Omega^1$ , la formulation variationnelle Lagrangienne totale est obtenue directement de 14.27 et dorénavant écrite sous la forme :

$$\begin{aligned} \int_{\Omega^0} \mathbf{S}(\mathbf{C}_{0 \rightarrow 1}) : (\mathbf{F}_{0 \rightarrow 1}^\top \cdot \nabla_{X_0} \mathbf{w}) \, dX_0 &= \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA_0 + \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}_{0 \rightarrow 1}^{-\top} \cdot \mathbf{N}_0) J_{0 \rightarrow 1} \, dA_0 \\ &+ \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dX_0 \end{aligned} \quad (14.37)$$

Les intégrales sont effectuées sur la géométrie  $\Omega_0$  de normale  $\mathbf{N}_0$  et le jacobien de la transformation est noté  $J_{0 \rightarrow 1}$ . On remarque ici que le tenseur de Piola-Kirchoff  $\mathbf{S}(\mathbf{C}_{0 \rightarrow 1})$  est évalué en fonction de  $\mathbf{C}_{0 \rightarrow 1} = \mathbf{F}_{0 \rightarrow 1}^\top \cdot \mathbf{F}_{0 \rightarrow 1}$  c.-à-d. à partir d'une déformation par rapport à la géométrie initiale. Il en sera de même pour le tenseur  $\mathbf{C}(\mathbf{C}_{0 \rightarrow 1})$  (d'ordre 4) provenant de la linéarisation.

Nous supposons donc que nous sommes pas en mesure de résoudre (en formulation Lagrangienne totale habituelle) ce problème et donc d'en déduire le déplacement  $\mathbf{u}_{0 \rightarrow 1}$  ainsi que  $\mathbf{F}_{0 \rightarrow 1}$ . Pour passer de la configuration  $\Omega^0$  à  $\Omega^2$ , la formulation variationnelle Lagrangienne totale serait aussi obtenue de 14.27 :

$$\begin{aligned} \int_{\Omega^0} \mathbf{S}(\mathbf{C}_{0 \rightarrow 2}) : (\mathbf{F}_{0 \rightarrow 2}^\top \cdot \nabla_{X_0} \mathbf{w}) \, dX_0 &= \int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA_0 + \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}_{0 \rightarrow 2}^{-\top} \cdot \mathbf{N}_0) J_{0 \rightarrow 2} \, dA_0 \\ &+ \int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dX_0 \end{aligned} \quad (14.38)$$

Le tenseur de Piola-Kirchoff  $\mathbf{S}(\mathbf{C}_{0 \rightarrow 2})$  est toujours évalué en fonction de  $\mathbf{C}_{0 \rightarrow 2} = \mathbf{F}_{0 \rightarrow 2}^\top \cdot \mathbf{F}_{0 \rightarrow 2}$  c.-à-d. encore à partir d'une déformation par rapport à la géométrie initiale. Puisque nous ne souhaitons pas résoudre directement ce problème, l'idée de base est de transformer cette formulation variationnelle sur la configuration  $\Omega^1$  déjà calculée. Il suffit alors de regarder chacun des termes de la formulation 14.38. Notons d'abord que :

$$\mathbf{C}_{0 \rightarrow 2} = \mathbf{F}_{0 \rightarrow 2}^\top \cdot \mathbf{F}_{0 \rightarrow 2} = (\mathbf{F}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1})^\top (\mathbf{F}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1}) = \mathbf{F}_{0 \rightarrow 1}^\top \cdot \mathbf{C}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1}$$

et donc que :

$$\mathbf{S}(\mathbf{C}_{0 \rightarrow 2}) = \mathbf{S}(\mathbf{F}_{0 \rightarrow 1}^\top \cdot \mathbf{C}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1})$$

Notons de plus que de l'équation 14.4, on a :

$$\nabla_{X_i} \mathbf{w} = \nabla_{X_j} \mathbf{w} \cdot \mathbf{F}_{i \rightarrow j}$$

### Le terme principal

On a alors :

$$\begin{aligned} \int_{\Omega^0} \mathbf{S}(\mathbf{C}_{0 \rightarrow 2}) : \left( \mathbf{F}_{0 \rightarrow 2}^\top \cdot \nabla_{X_0} \mathbf{w} \right) dX_0 &= \int_{\Omega^1} \mathbf{S}(\mathbf{C}_{0 \rightarrow 2}) : \left( (\mathbf{F}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1})^\top \cdot (\nabla_{X_1} \mathbf{w} \cdot \mathbf{F}_{0 \rightarrow 1}) \right) J_{0 \rightarrow 1}^{-1} dX_1 \\ &= \int_{\Omega^1} \tilde{\mathbf{S}}(\mathbf{C}_{0 \rightarrow 2}) : \left( \mathbf{F}_{1 \rightarrow 2}^\top \cdot \nabla_{X_1} \mathbf{w} \right) dX_1 \end{aligned}$$

où l'on a posé :

$$\tilde{\mathbf{S}}(\mathbf{C}_{0 \rightarrow 2}) = J_{0 \rightarrow 1}^{-1} \mathbf{F}_{0 \rightarrow 1} \cdot \mathbf{S}(\mathbf{C}_{0 \rightarrow 2}) \cdot \mathbf{F}_{0 \rightarrow 1}^\top \quad (14.39)$$

et où on a utilisé les propriétés des produits tensoriels (voir les exercices de la série B.3). C'est le terme de base de la formulation Lagrangienne actualisée. Notons de plus que le tenseur de Cauchy s'écrit :

$$\boldsymbol{\sigma} = \frac{1}{J_{0 \rightarrow 2}} \mathbf{F}_{0 \rightarrow 2} \cdot \mathbf{S}(\mathbf{C}_{0 \rightarrow 2}) \cdot \mathbf{F}_{0 \rightarrow 2}^\top = \frac{1}{J_{1 \rightarrow 2}} \mathbf{F}_{1 \rightarrow 2} \cdot \tilde{\mathbf{S}} \cdot \mathbf{F}_{1 \rightarrow 2}^\top \quad (14.40)$$

Voici quelques remarques supplémentaires :

1. Le tenseur  $\mathbf{F}_{0 \rightarrow 1}$  est supposé connu (calculé préalablement). Le déplacement  $\mathbf{u}_{0 \rightarrow 1}$  n'intervient explicitement nulle part ;
2. Les inconnues du problème sont  $\mathbf{u}_{1 \rightarrow 2}$  et par conséquent  $\mathbf{F}_{1 \rightarrow 2}$  et  $\mathbf{C}_{1 \rightarrow 2}$  ;
3. Pour évaluer  $\mathbf{C}_{0 \rightarrow 2}$ , on utilise l'expression  $(\mathbf{F}_{0 \rightarrow 1}^\top \cdot \mathbf{C}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1})$  en supposant le tenseur  $\mathbf{F}_{0 \rightarrow 1}$  connu ;
4. L'expression du tenseur de Piola-Kirchoff sur  $\Omega^1$  compose les déformations antérieures et est donc un peu plus complexe. La formulation Lagrangienne totale correspond au cas où  $\mathbf{F}_{0 \rightarrow 1} = \mathbf{I}$  c.-à-d. qu'il n'y a aucune déformation préalable.

On doit effectuer le même travail pour les termes de linéarisation. Ainsi les deux premiers termes de 14.31 deviennent, en se servant des propriétés des produits tensoriels des exercices de la série B.3 :

$$\begin{aligned} \int_{\Omega^0} \mathbf{S}(\mathbf{C}_{0 \rightarrow 2}) : \left( \nabla_{X_0}^\top (\boldsymbol{\delta}_u) \cdot \nabla_{X_0} \mathbf{w} \right) dX_0 \\ = \int_{\Omega^1} J_{0 \rightarrow 1}^{-1} \mathbf{F}_{0 \rightarrow 1} \cdot \mathbf{S}(\mathbf{F}_{0 \rightarrow 1}^\top \cdot \mathbf{C}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1}) \mathbf{F}_{0 \rightarrow 1}^\top : \left( \nabla_{X_1}^\top (\boldsymbol{\delta}_u) \cdot \nabla_{X_1} \mathbf{w} \right) dX_1 \quad (14.41) \\ = \int_{\Omega^1} \tilde{\mathbf{S}}(\mathbf{C}_{0 \rightarrow 2}) : \left( \nabla_{X_1}^\top (\boldsymbol{\delta}_u) \cdot \nabla_{X_1} \mathbf{w} \right) dX_1 \end{aligned}$$

et de même :

$$\begin{aligned} \int_{\Omega^0} \left( \mathbf{c}(\mathbf{C}_{0 \rightarrow 2}) : \left( \mathbf{F}_{0 \rightarrow 2}^\top \cdot \nabla_{X_0} (\boldsymbol{\delta}_u) \right) \right) : \left( \mathbf{F}_{0 \rightarrow 2}^\top \cdot \nabla_{X_0} \mathbf{w} \right) dX_0 = \\ \int_{\Omega^1} \left( \tilde{\mathbf{c}}(\mathbf{C}_{0 \rightarrow 2}) : \left( \mathbf{F}_{1 \rightarrow 2}^\top \cdot \nabla_{X_1} (\boldsymbol{\delta}_u) \right) \right) : \left( \mathbf{F}_{1 \rightarrow 2}^\top \cdot \nabla_{X_1} \mathbf{w} \right) dX_1 \end{aligned} \quad (14.42)$$

où le tenseur  $\tilde{\mathbf{C}}(\mathbf{C}_{0 \rightarrow 2})$  a pour composantes :

$$(\tilde{\mathbf{C}}(\mathbf{C}_{0 \rightarrow 2}))_{mnop} = J_{0 \rightarrow 1}^{-1} (\mathbf{C}(\mathbf{C}_{0 \rightarrow 2}))_{ijkl} (\mathbf{F}_{0 \rightarrow 1})_{mi} (\mathbf{F}_{0 \rightarrow 1})_{nj} (\mathbf{F}_{0 \rightarrow 1})_{ok} (\mathbf{F}_{0 \rightarrow 1})_{pl} \quad (14.43)$$

Nous pouvons maintenant procéder ainsi pour tous les autres termes de la formulation variationnelle.

### Le terme source

$$\int_{\Omega^0} \mathbf{r}_0 \cdot \mathbf{w} \, dX_0 = \int_{\Omega^1} \mathbf{r}_0 \cdot \mathbf{w} J_{0 \rightarrow 1}^{-1} \, dX_1 = \int_{\Omega^1} \mathbf{r}_1 \cdot \mathbf{w} \, dX_1 \text{ où } \mathbf{r}_1 = J_{0 \rightarrow 1}^{-1} \mathbf{r}_0$$

### La condition de Neumann

$$\int_{\Gamma_N^0} \mathbf{h}_0 \cdot \mathbf{w} \, dA_0 = \int_{\Gamma_N^1} \mathbf{h}_0 \cdot \mathbf{w} (J_s)_{0 \rightarrow 1}^{-1} \, dA_1 = \int_{\Gamma_N^1} \mathbf{h}_1 \cdot \mathbf{w} \, dA_1 \text{ où } \mathbf{h}_1 = (J_s)_{0 \rightarrow 1}^{-1} \mathbf{h}_0$$

### La pression suiveuse

Le jacobien surfacique se transforme suivant la formule :

$$(J_s)_{0 \rightarrow 2} = \frac{dA_2}{dA_0} = \frac{dA_2}{dA_1} \frac{dA_1}{dA_0} = (J_s)_{1 \rightarrow 2} (J_s)_{0 \rightarrow 1}$$

et le vecteur normal suivant la formule de transformation de la normale 14.11 :

$$\mathbf{N}_1 = \frac{J_{0 \rightarrow 1}}{(J_s)_{0 \rightarrow 1}} \mathbf{F}_{0 \rightarrow 1}^{-\top} \cdot \mathbf{N}_0$$

et il s'en suit que :

$$\begin{aligned} \int_{\Gamma_P^0} P \mathbf{w} \cdot (\mathbf{F}_{0 \rightarrow 2}^{-\top} \cdot \mathbf{N}_0) J_{0 \rightarrow 2} \, dA_0 &= \int_{\Gamma_P^1} P \mathbf{w} \cdot \left( (\mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot \mathbf{F}_{0 \rightarrow 1}^{-\top}) \cdot \mathbf{N}_0 \right) J_{0 \rightarrow 1} J_{1 \rightarrow 2} (J_s)_{0 \rightarrow 1}^{-1} \, dA_1 \\ &= \int_{\Gamma_P^1} P \mathbf{w} \cdot (\mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot \mathbf{N}_1) J_{1 \rightarrow 2} \, dA_1 \end{aligned}$$

En ce qui concerne la linéarisation de ce terme, on doit transformer sur  $\Omega^1$  l'expression :

$$\int_{\Gamma_P^0} P \mathbf{w} \cdot \left[ \left( \mathbf{F}_{0 \rightarrow 2}^{-\top} : \nabla_{X_0} \boldsymbol{\delta}_u \right) (\mathbf{F}_{0 \rightarrow 2}^{-\top} \cdot \mathbf{N}_0) - \left( \mathbf{F}_{0 \rightarrow 2}^{-\top} \cdot (\nabla_{X_0} \boldsymbol{\delta}_u)^\top \cdot \mathbf{F}_{0 \rightarrow 2}^{-\top} \right) \cdot \mathbf{N}_0 \right] J_{0 \rightarrow 2} \, dA_0$$

Le premier terme devient :

$$\begin{aligned}
& \int_{\Gamma_P^0} P\mathbf{w} \cdot \left( \mathbf{F}_{0 \rightarrow 2}^{-\top} : \nabla_{X_0} \boldsymbol{\delta}_u \right) (\mathbf{F}_{0 \rightarrow 2}^{-\top} \cdot \mathbf{N}_0) J_{0 \rightarrow 2} dA_0 \\
&= \int_{\Gamma_P^1} P\mathbf{w} \cdot \left( (\mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot \mathbf{F}_{0 \rightarrow 1}^{-\top}) : (\nabla_{X_1} \boldsymbol{\delta}_u \cdot \mathbf{F}_{0 \rightarrow 1}) \right) \cdot \left( \mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot \mathbf{F}_{0 \rightarrow 1}^{-\top} \cdot \mathbf{N}_0 \right) J_{1 \rightarrow 2} J_{0 \rightarrow 1} (J_s)_{0 \rightarrow 1}^{-1} dA_1 \\
&= \int_{\Gamma_P^1} P\mathbf{w} \cdot \left[ \mathbf{F}_{1 \rightarrow 2}^{-\top} : (\nabla_{X_1} \boldsymbol{\delta}_u) \right] \cdot (\mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot \mathbf{N}_1) J_{1 \rightarrow 2} dA_1
\end{aligned}$$

L'autre terme de la linéarisation se traite de manière similaire et on a :

$$\begin{aligned}
& - \int_{\Gamma_P^0} P\mathbf{w} \cdot \left[ \left( \mathbf{F}_{0 \rightarrow 2}^{-\top} \cdot (\nabla_{X_0} \boldsymbol{\delta}_u)^\top \cdot \mathbf{F}_{0 \rightarrow 2}^{-\top} \right) \cdot \mathbf{N}_0 \right] J_{0 \rightarrow 2} dA_0 \\
&= - \int_{\Gamma_P^1} P\mathbf{w} \cdot \left[ \left( \mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot (\nabla_{X_1} \boldsymbol{\delta}_u)^\top \cdot \mathbf{F}_{1 \rightarrow 2}^{-\top} \right) \cdot \mathbf{N}_1 \right] J_{1 \rightarrow 2} dA_1
\end{aligned}$$

### Formulation Lagrangienne actualisée linéarisée (en déplacement)

La formulation variationnelle Lagrangienne actualisée complète s'écrit donc :

$$\begin{aligned}
& \int_{\Omega^1} \tilde{\mathbf{S}}(\mathbf{C}_{0 \rightarrow 2}) : \left( \nabla_{X_1}^\top (\boldsymbol{\delta}_u) \cdot \nabla_{X_1} \mathbf{w} \right) dX_1 + \int_{\Omega^1} \tilde{\mathbf{C}}(\mathbf{C}_{0 \rightarrow 2}) : \left( \mathbf{F}_{1 \rightarrow 2}^\top \cdot \nabla_{X_1} (\boldsymbol{\delta}_u) \right) : \left( \mathbf{F}_{1 \rightarrow 2}^\top \cdot \nabla_{X_1} \mathbf{w} \right) dX_1 \\
& - \int_{\Gamma_P^1} P\mathbf{w} \cdot \left[ \mathbf{F}_{1 \rightarrow 2}^{-\top} : (\nabla_{X_1} \boldsymbol{\delta}_u) \right] \cdot (\mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot \mathbf{N}_1) J_{1 \rightarrow 2} dA_1 \\
& + \int_{\Gamma_P^1} P\mathbf{w} \cdot \left[ \left( \mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot (\nabla_{X_1} \boldsymbol{\delta}_u)^\top \cdot \mathbf{F}_{1 \rightarrow 2}^{-\top} \right) \cdot \mathbf{N}_1 \right] J_{1 \rightarrow 2} dA_1 \\
&= - \int_{\Omega^1} \tilde{\mathbf{S}}(\mathbf{C}_{0 \rightarrow 2}) : \left( \mathbf{F}_{1 \rightarrow 2}^\top \cdot \nabla_{X_1} \mathbf{w} \right) dX_1 + \int_{\Omega^1} \mathbf{r}_1 \cdot \mathbf{w} dX_1 + \int_{\Gamma_N^1} \mathbf{h}_1 \cdot \mathbf{w} dA_1 \\
& + \int_{\Gamma_P^1} P\mathbf{w} \cdot (\mathbf{F}_{1 \rightarrow 2}^{-\top} \cdot \mathbf{N}_1) J_{1 \rightarrow 2} dA_1
\end{aligned} \tag{14.44}$$

### 14.8.2 Formulation Lagrangienne actualisée mixte

Si on se réfère à la formulation variationnelle mixte 14.33 déjà obtenue, on montre de manière similaire à ce qui a été fait à la section 14.8.1, que la formulation variationnelle complète s'écrit :

$$\begin{aligned}
R_1 &= \int_{\Omega^1} \tilde{\mathbf{S}}(\mathbf{C}_{0 \rightarrow 2}) : (\mathbf{F}_{1 \rightarrow 2}^\top \cdot \nabla_{X_1} \mathbf{w}) dX_1 - \int_{\Gamma_N^1} \mathbf{h}_1 \cdot \mathbf{w} dA_1 - \int_{\Omega^1} \mathbf{r}_1 \cdot \mathbf{w} dX_1 \\
&= \int_{\Omega^1} \tilde{\mathbf{S}}'(\mathbf{C}_{0 \rightarrow 2}) : (\mathbf{F}_{1 \rightarrow 2}^\top \cdot \nabla_{X_1} \mathbf{w}) dV - \int_{\Omega^1} p J_{1 \rightarrow 2} \mathbf{F}_{1 \rightarrow 2}^{-\top} : \nabla_{X_1} \mathbf{w} dX_1 \\
&\quad - \int_{\Gamma_N^1} \mathbf{h}_1 \cdot \mathbf{w} dA_1 - \int_{\Omega^1} \mathbf{r}_1 \cdot \mathbf{w} dX_1 \\
R_2 &= \int_{\Omega^1} \left( -(J_{1 \rightarrow 2} - J_{0 \rightarrow 1}^{-1}) - \frac{1}{k} p J_{0 \rightarrow 1}^{-1} \right) q dV
\end{aligned}$$

où  $\tilde{\mathbf{S}}'(\mathbf{C}_{0 \rightarrow 2}) = J_{0 \rightarrow 1}^{-1} \mathbf{F}_{0 \rightarrow 1} \cdot \mathbf{S}'(\mathbf{C}_{0 \rightarrow 2}) \cdot \mathbf{F}_{0 \rightarrow 1}^\top$ ,  $\mathbf{r}_1 = J_{0 \rightarrow 1}^{-1} \mathbf{r}_0$  et  $\mathbf{h}_1 = (J_s)_{0 \rightarrow 1}^{-1} \mathbf{h}_0$ . La formulation variationnelle complète s'écrit :

$$\left\{ \begin{array}{l} \frac{\partial R_1}{\partial \mathbf{u}} \cdot \delta \mathbf{u} - \int_{\Omega^1} \delta_p J_{1 \rightarrow 2} \mathbf{F}_{1 \rightarrow 2}^{-\top} : \nabla_{X_1} \mathbf{w} dX_1 = -R_1 \\ - \int_{\Omega^1} q J_{1 \rightarrow 2} \mathbf{F}_{1 \rightarrow 2}^{-\top} : \nabla_{X_1} \delta \mathbf{u} dX_1 - \int_{\Omega^1} \frac{1}{k} \delta_p q J_{0 \rightarrow 1}^{-1} dX_1 = -R_2 \end{array} \right. \quad (14.45)$$

où le terme  $\frac{\partial R_1}{\partial \mathbf{u}} \cdot \delta \mathbf{u}$  prend exactement la même forme que pour la formulation en déplacement seulement (voir les équations 14.41 et 14.42).

### 14.8.3 Équivalence des formulations

En absence de remaillage, les formulations Lagrangiennes totale et actualisée sont équivalentes. En effet ...

### 14.8.4 Algorithme complet

#### Cumul des prédéformations

Supposons maintenant que nous en sommes à l'étape  $i$  c.-à-d. nous avons déjà calculé  $i$  configurations  $\Omega^i$  et que nous espérons obtenir  $\Omega^{i+1}$ . Le même raisonnement que précédemment peut encore être appliqué, en remplaçant  $F_{0 \rightarrow 2}$  par  $\mathbf{F}_{0 \rightarrow i+1}$ . Il suffit alors de constater que :

$$\mathbf{F}_{0 \rightarrow i+1} = \mathbf{F}_{i \rightarrow i+1} \cdot \mathbf{F}_{i-1 \rightarrow i} \cdot \mathbf{F}_{i-2 \rightarrow i-1} \cdot \dots \cdot \mathbf{F}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1}$$

On peut ainsi en déduire  $\mathbf{C}_{0 \rightarrow i+1}$ , qui nous permettra de calculer les tenseurs de Piola-Kirchoff et d'élasticité. Comme les géométries intermédiaires sont détruites au fur et à mesure, il suffit de cumuler ces produits dans un tenseur, toujours noté  $\mathbf{F}_{0 \rightarrow 1}$  dans l'algorithme qui suit.

**Algorithme 14.1 : Formulation Lagrangienne actualisée**

0. Initialisation :  $\Omega^1 = \Omega^0$  et donc  $\mathbf{F}_{0 \rightarrow 1} = \mathbf{I}$ ,  $J_{0 \rightarrow 1} = 1$  et  $(J_s)_{0 \rightarrow 1} = 1$  ;
1. Résolution du système 14.44 pour obtenir  $\mathbf{u}_{1 \rightarrow 2}$  et  $\mathbf{F}_{1 \rightarrow 2}$  ;
  - 1.1  $\mathbf{F}_{0 \rightarrow 1}$  est supposé connu ;
  - 1.2 on calcule  $\mathbf{C}_{0 \rightarrow 2} = \mathbf{F}_{0 \rightarrow 1}^\top \cdot \mathbf{C}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1}$  ;
  - 1.3 on calcule  $\mathbf{S}(\mathbf{C}_{0 \rightarrow 2})$  et  $\mathcal{C}(\mathbf{C}_{0 \rightarrow 2})$  ;
  - 1.4 on en déduit les tenseurs  $\tilde{\mathbf{S}}$  et  $\tilde{\mathcal{C}}$  par les relations 14.39 et 14.43 ;
  - 1.4 on calcule (si nécessaire) le tenseur de Cauchy par la relation 14.40 ;
2. Mise à jour de la géométrie :  $\Omega^2 = \Omega^1 + \mathbf{u}_{1 \rightarrow 2}$  ;
  - 2.1  $\Omega^2$  devient  $\Omega^1$  ;
  - 2.2  $\mathbf{F}_{0 \rightarrow 1}$  est remplacé par  $\mathbf{F}_{1 \rightarrow 2} \cdot \mathbf{F}_{0 \rightarrow 1}$  (cumul des prédéformations). Par conséquent,  $J_{0 \rightarrow 1}$  est remplacé par  $J_{1 \rightarrow 2} J_{0 \rightarrow 1}$  et  $(J_s)_{0 \rightarrow 1}$  est remplacé par  $(J_s)_{1 \rightarrow 2} (J_s)_{0 \rightarrow 1}$  ;
3. Remaillage (si nécessaire) ;
4. Réinterpolation des différentes variables sur le nouveau maillage (si nécessaire) ;
5. Retour à l'étape 1 pour l'incrément suivant

▲

Les étapes 3 et 4 de remaillage et de réinterpolation sont délicates et feront l'objet d'une discussion détaillée un peu plus loin.

## 14.9 Exercices

1. Soit  $\boldsymbol{\sigma}$  et  $\boldsymbol{\tau}$  des tenseurs d'ordre 2 et supposons que  $\boldsymbol{\sigma}$  est symétrique. Montrer que :

$$\boldsymbol{\sigma} : \left( \frac{\boldsymbol{\tau} + \boldsymbol{\tau}^T}{2} \right) = \boldsymbol{\sigma} : \boldsymbol{\tau}$$

2. Soit  $\mathbf{C}$  un tenseur symétrique d'ordre 4 et  $\boldsymbol{\tau}$  un tenseur d'ordre 2. Montrer que :

$$\mathbf{C} : \left( \frac{\boldsymbol{\tau} + \boldsymbol{\tau}^T}{2} \right) = \mathbf{C} : \boldsymbol{\tau}$$

3. En utilisant la relation 14.1, montrer que :

$$\mathbf{F}(\mathbf{u} + \mathbf{w}) = \mathbf{F}(\mathbf{u}) + \nabla_X(\mathbf{w})$$

4. Montrer que :

$$\frac{\partial(\mathbf{B} : \mathbf{B})}{\partial \mathbf{B}} = \mathbf{B}$$

5. Montrer que dans le cas des petites déformations, la condition :

$$\det \mathbf{F} \simeq 1$$

est équivalente à :

$$\nabla \cdot \mathbf{u} \simeq 0$$

6. Montrer à l'aide des expressions (B.20) que :

$$\text{tr} \left( \frac{\partial J_1}{\partial \mathbf{C}} \cdot \mathbf{C} \right) = \text{tr} \left( \frac{\partial J_2}{\partial \mathbf{C}} \cdot \mathbf{C} \right) = 0$$



## Chapitre 15

# Problèmes de contact unilatéral et frottant

### 15.1 Introduction

Le traitement des problèmes de contact est crucial dans bon nombre d'applications. Un exemple classique est celui d'un pneu se déformant au contact de la route sous le poids d'un véhicule. Le frottement du pneu sur la route est à la fois bénéfique, car il empêche le véhicule de dérapier dans les courbes, et nuisible car le frottement fait augmenter la consommation d'essence. Il faut donc faire un compromis et pour y arriver, il faut bien comprendre les phénomènes liés au contact frottant.

Nous allons chercher à modéliser la déformation de deux (et même plusieurs) corps qui entrent en contact et se déforment mutuellement. Cela inclut bien entendu le cas d'un objet déformable  $\Omega$  entrant en contact avec un corps non déformable (rigide), cas un peu plus simple. Dans le cas où les deux corps sont déformables, il est fréquent de choisir l'un d'entre eux comme le maître (noté  $\Omega^R$ ) et l'autre comme l'esclave (noté simplement  $\Omega$ ). Le choix du maître et de l'esclave est *a priori* arbitraire, mais a des conséquences sur les résultats numériques. Dans le cas où l'un des deux corps est non déformable (rigide), ce sera le maître.

#### 15.1.1 Distance signée

La notion de distance entre deux corps susceptibles d'entrer en contact est fondamentale et mérite de s'y attarder un moment. On introduit dans un premier temps la fonction distance au corps rigide  $\Omega^R$  :

$$d^{\Omega^R}(\mathbf{x}) = \inf_{\mathbf{y} \in \Omega^R} \|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} - \mathbf{p}^{\Omega^R}(\mathbf{x})\|$$

qui est une fonction positive ou nulle. La notion de distance est utile mais celle de distance signée l'est encore plus. On la définit comme la distance à  $\Omega^R$  moins la distance à son complémentaire ou encore :

$$b^{\Omega^R}(\mathbf{x}) = d^{\Omega^R}(\mathbf{x}) - d^{\mathcal{R}^n \setminus \Omega^R}(\mathbf{x})$$

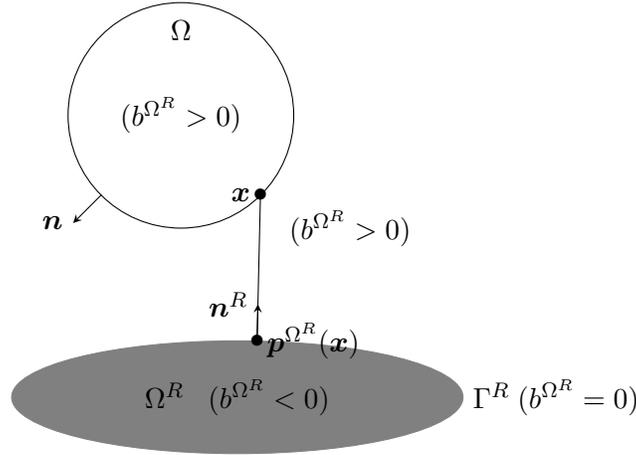


Figure 15.1 – Fonction distance signée (par projection)

où  $\mathcal{R}^n \setminus \Omega^R$  est l'ensemble complémentaire de  $\Omega^R$ . On notera  $\mathbf{p}^{\Omega^R}(\mathbf{x})$  le projeté de  $\mathbf{x}$  sur la frontière de  $\Omega^R$  dont le vecteur normal unitaire est noté  $\mathbf{n}^R$ . La fonction  $b^{\Omega^R}$  est donc positive à l'extérieur de  $\Omega^R$ , négative à l'intérieur et nulle sur sa frontière  $\Gamma^R$  (voir la figure 15.1). La fonction distance signée possède plusieurs propriétés qui nous seront utiles par la suite et qui nous allons maintenant établir.

♣ **Lemme 15.16**

Le vecteur  $\mathbf{x} - \mathbf{p}^{\Omega^R}(\mathbf{x})$  est orthogonal à  $\Gamma^R$ . De plus,  $\|\nabla b^{\Omega^R}\| = 1$  de sorte que  $\nabla b^{\Omega^R} = \mathbf{n}^R$ .

**Démonstration :**

Soit  $\mathbf{x}$  un point de  $\partial\Omega$  et on considère une paramétrisation  $\mathbf{x}^R(\xi_1, \xi_2)$  de la surface maître  $\Gamma^R$ . On cherche le point de cette surface le plus près de  $\mathbf{x}$  c.-à-d. qui minimise la distance entre  $\mathbf{x}$  et  $\mathbf{x}^R$  ou encore mieux, qui minimise le carré de cette distance. Ce minimum est caractérisé par :

$$\frac{\partial}{\partial \xi_k} \frac{1}{2} \|\mathbf{x} - \mathbf{x}^R(\xi_1, \xi_2)\|^2 = \frac{\partial}{\partial \xi_k} (\mathbf{x} - \mathbf{x}^R(\xi_1, \xi_2)) \cdot (\mathbf{x} - \mathbf{x}^R(\xi_1, \xi_2)) = 0 \text{ pour } k = 1, 2$$

ou plus précisément :

$$(\mathbf{x} - \mathbf{x}^R) \cdot \frac{\partial \mathbf{x}^R}{\partial \xi_k} = 0 \text{ pour } k = 1, 2 \quad (15.1)$$

et le résultat s'en suit car les vecteurs  $\frac{\partial \mathbf{x}^R}{\partial \xi_k}$  sont tangents à la surface.

On notera  $\mathbf{p}^{\Omega^R}(\mathbf{x})$  le point optimal de sorte que le vecteur  $(\mathbf{x} - \mathbf{p}^{\Omega^R}(\mathbf{x}))$  soit orthogonal à la frontière de  $\Omega^R$  et que :

$$|b^{\Omega^R}(\mathbf{x})| = \|\mathbf{x} - \mathbf{p}^{\Omega^R}(\mathbf{x})\|$$

Le vecteur  $(\mathbf{x} - \mathbf{p}^{\Omega^R}(\mathbf{x}))$  pointe vers l'extérieur de  $\Omega^R$  si  $\mathbf{x}$  est lui-même à l'extérieur de  $\Omega^R$  et vers l'intérieur dans le cas contraire. Par conséquent, on peut écrire que :

$$\mathbf{n}^R = \frac{\mathbf{x} - \mathbf{p}^{\Omega^R}(\mathbf{x})}{b^{\Omega^R}(\mathbf{x})}$$

qui est toujours dirigé vers l'extérieur de  $\Omega^R$ . Enfin, puisque  $\Gamma^R$  est la surface de niveau 0 de la fonction  $b^{\Omega^R}$ , on a aussi :

$$\mathbf{n}^R = \frac{\nabla b^{\Omega^R}}{\|\nabla b^{\Omega^R}\|}$$

Par définition, le taux d'accroissement de la fonction distance dans la direction normale à la frontière de  $\Omega^R$  est 1 c.-à-d.

$$1 = \frac{\partial b^{\Omega^R}(\mathbf{x})}{\partial \mathbf{n}^R} = \nabla b^{\Omega^R} \cdot \mathbf{n}^R = \nabla b^{\Omega^R} \cdot \frac{\nabla b^{\Omega^R}}{\|\nabla b^{\Omega^R}\|} = \|\nabla b^{\Omega^R}\|$$

de sorte que :

$$\mathbf{n}^R = \frac{\mathbf{x} - \mathbf{p}^{\Omega^R}(\mathbf{x})}{b^{\Omega^R}(\mathbf{x})} = \nabla b^{\Omega^R}$$

De cette dernière relation, on déduit immédiatement que :

$$\mathbf{p}^{\Omega^R}(\mathbf{x}) = \mathbf{x} - b^{\Omega^R}(\mathbf{x}) \nabla b^{\Omega^R}$$

d'où :

$$b^{\Omega^R}(\mathbf{x}) = (\mathbf{x} - \mathbf{p}^{\Omega^R}(\mathbf{x})) \cdot \mathbf{n}^R \tag{15.2}$$



◆ **Remarque 15.1**

La démonstration précédente nous indique aussi comment calculer le point de projection. Nous verrons comment établir la paramétrisation de la surface un peu plus loin. Ayant cette paramétrisation, il suffit de résoudre le système 15.1 pour  $\xi_1$  et  $\xi_2$ . Une méthode de Newton classique permettra d'y arriver facilement. La matrice jacobienne associée à ce problème est tout simplement :

$$J(\xi_1, \xi_2) = \begin{bmatrix} \frac{\partial \mathbf{x}^R}{\partial \xi_1} \cdot \frac{\partial \mathbf{x}^R}{\partial \xi_1} - \frac{\partial^2 \mathbf{x}^R}{\partial \xi_1^2} \cdot (\mathbf{x} - \mathbf{x}^R) & \frac{\partial \mathbf{x}^R}{\partial \xi_1} \cdot \frac{\partial \mathbf{x}^R}{\partial \xi_2} - \frac{\partial^2 \mathbf{x}^R}{\partial \xi_1 \partial \xi_2} \cdot (\mathbf{x} - \mathbf{x}^R) \\ \frac{\partial \mathbf{x}^R}{\partial \xi_1} \cdot \frac{\partial \mathbf{x}^R}{\partial \xi_2} - \frac{\partial^2 \mathbf{x}^R}{\partial \xi_1 \partial \xi_2} \cdot (\mathbf{x} - \mathbf{x}^R) & \frac{\partial \mathbf{x}^R}{\partial \xi_2} \cdot \frac{\partial \mathbf{x}^R}{\partial \xi_2} - \frac{\partial^2 \mathbf{x}^R}{\partial \xi_2^2} \cdot (\mathbf{x} - \mathbf{x}^R) \end{bmatrix}$$



## 15.2 Contact unilatéral en petites déformations

Nous considérons dans un premier temps un problème en élasticité linéaire (petites déformations) dont nous avons préalablement discuté au chapitre 11. Rappelons que dans ce cas,  $\Omega^0 = \Omega^t = \Omega$  c.-à-d. que les déformations sont suffisamment petites pour confondre la géométrie initiale et la géométrie déformée. On aura donc  $\mathbf{X} = \mathbf{x}$  et dans ce cas relativement simple, nous faisons implicitement l'hypothèse suivante :

**H1** : Dans ce cadre des petites déformations, le vecteur normal  $\mathbf{n}^R$  ainsi que le point de contact  $p^{\Omega^R}(\mathbf{x})$  ne dépendent pas du déplacement  $\mathbf{u}$ . À toute fin pratique, cela suppose que les corps sont initialement en contact et donc que  $\mathbf{n} = -\mathbf{n}^R$ .

L'hypothèse H1 signifie en particulier que  $p^{\Omega^R}(\mathbf{x}) = p^{\Omega^R}(\mathbf{x} + \mathbf{u})$  et que  $\mathbf{n}^R(\mathbf{x} + \mathbf{u}) = \mathbf{n}^R(\mathbf{x})$ . On définit maintenant la fonction d'écart entre les deux domaines.<sup>1</sup>

$$\begin{aligned} g(\mathbf{x}) &= (\mathbf{x} + \mathbf{u} - P^{\Omega^R}(\mathbf{x} + \mathbf{u})) \cdot \mathbf{n}^R(\mathbf{x} + \mathbf{u}) \\ &= (\mathbf{x} + \mathbf{u} - P^{\Omega^R}(\mathbf{x})) \cdot \mathbf{n}^R(\mathbf{x}) \quad (\text{hypothèse H1}) \\ &= (\mathbf{x} - P^{\Omega^R}(\mathbf{x})) \cdot \mathbf{n}^R(\mathbf{x}) + \mathbf{u} \cdot \mathbf{n}^R(\mathbf{x}) \end{aligned}$$

c'est-à-dire :

$$g(\mathbf{x}) = g_0(\mathbf{x}) + \mathbf{u} \cdot \mathbf{n}^R(\mathbf{x}) \quad \text{où } g_0(\mathbf{x}) = (\mathbf{x} - P^{\Omega^R}(\mathbf{x})) \cdot \mathbf{n}^R(\mathbf{x}) \quad (15.3)$$

Le terme  $g_0(\mathbf{x})$  est appelé l'écart initial et ne dépend pas du déplacement  $\mathbf{u}$ . On a de plus que :

$$\frac{\partial p^{\Omega^R}(\mathbf{x} + \mathbf{u})}{\partial \mathbf{u}} \cdot \delta \mathbf{u} = \frac{\partial p^{\Omega^R}(\mathbf{x})}{\partial \mathbf{u}} \cdot \delta \mathbf{u} = 0 \quad \text{ainsi que} \quad \frac{\partial \mathbf{n}^R(\mathbf{x} + \mathbf{u})}{\partial \mathbf{u}} \cdot \delta \mathbf{u} = \frac{\partial \mathbf{n}^R(\mathbf{x})}{\partial \mathbf{u}} \cdot \delta \mathbf{u} = 0$$

d'où l'on tire le lemme suivant.

### ♣ Lemme 15.17

On a la propriété suivante :

$$\frac{\partial g(\mathbf{x})}{\partial \mathbf{u}} \cdot \delta \mathbf{u} = \mathbf{n}^R \cdot \delta \mathbf{u} \quad (15.4)$$

♣

Dans un premier temps, le problème peut s'écrire sous une forme classique :

$$\begin{cases} -\nabla \cdot \boldsymbol{\sigma} = \mathbf{r} & \text{dans } \Omega \\ \mathbf{u} = \mathbf{u}_g & \text{sur } \Gamma_D \\ \boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{h} & \text{sur } \Gamma_N \end{cases}$$

1. La fonction d'écart n'est rien d'autres que le « gap » en anglais d'où la notation  $g$ .

En ce qui concerne le contact, nous supposons, dans un premier temps, qu'il n'y a pas de force tangentielle (liée au frottement) et donc que  $\boldsymbol{\sigma}_T = 0$  sur la frontière  $\Gamma_C$  susceptible d'entrer en contact. Nous traiterons donc d'abord la condition de non-pénétration sur la frontière en contact. On doit alors imposer sur  $\Gamma_C$  des conditions de type Kuhn-Tucker de la forme :

$$\begin{cases} \sigma_{nn} & \leq 0 \\ b^{\Omega^R}(\mathbf{u}) & \geq 0 \\ b^{\Omega^R}(\mathbf{u})\sigma_{nn} & = 0 \end{cases}$$

Cela signifie que s'il y a contact ( $b^{\Omega^R} = 0$ ), alors le corps rigide agit en compression ( $\sigma_{nn} \leq 0$ ) sur le corps déformable. Inversement, s'il n'y a pas contact ( $b^{\Omega^R}(\mathbf{u}) > 0$ ), alors la trace normale de la traction est nulle ( $\sigma_{nn} = 0$ ). La loi de contact unilatéral est illustrée à la figure 15.2<sup>2</sup>. L'imposition de telles conditions n'est certainement pas triviale et de nombreuses méthodes ont été introduites pour y arriver.

Rappelons du chapitre 11 que la formulation variationnelle est de la forme :

$$\int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\gamma}(\mathbf{w}) \, dv - \int_{\Gamma} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \mathbf{w} \, ds = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv$$

où  $\boldsymbol{\sigma} = \mathcal{C} : \boldsymbol{\gamma}(\mathbf{u})$ . Puisque nous avons des conditions de Dirichlet  $\mathbf{u} = \mathbf{u}_g$  sur  $\Gamma_D$ , alors les fonctions tests s'y annulent et  $\mathbf{w} = 0$ . En posant  $\boldsymbol{\sigma} \cdot \mathbf{n} = \sigma_{nn}\mathbf{n} + \boldsymbol{\sigma}_T$  et  $\mathbf{w} = (\mathbf{w} \cdot \mathbf{n})\mathbf{n} + \mathbf{w}_T$ , on trouve :

$$\int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\gamma}(\mathbf{w}) \, dv - \int_{\Gamma_C} \sigma_{nn}\mathbf{w} \cdot \mathbf{n} \, ds = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv + \int_{\Gamma_N} \mathbf{h} \cdot \mathbf{w} \, ds$$

Si on réécrit le tout en correction, on trouve :

$$\int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\gamma}(\mathbf{w}) \, dv - \int_{\Gamma_C} \sigma_{nn}\mathbf{w} \cdot \mathbf{n} \, ds = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} \, dv + \int_{\Gamma_N} \mathbf{h} \cdot \mathbf{w} \, ds$$

Le contact unilatéral conduit, comme nous le verrons, non pas aux égalités variationnelles auxquelles nous sommes maintenant habitués, mais bien à des inéquations variationnelles.

Sous l'hypothèse H1, la distance entre le corps rigide est aussi donnée par :

$$b^{\Omega^R}(\mathbf{x}) = b_0^{\Omega^R}(\mathbf{x}) - \boldsymbol{\delta}_u \cdot \mathbf{n} \geq 0$$

et on définit les déplacements admissibles comme étant ceux qui respectent la condition de non-pénétration. On pose alors :

$$\mathcal{K} = \{v \in H_{\Gamma_0}^1(\Omega) \mid \mathbf{v} \cdot \mathbf{n} \leq b_0^{\Omega^R}(\mathbf{x})\}$$

---

2. Le lecteur remarquera peut-être que nous utilisons une convention différente de ce que l'on retrouve habituellement dans la littérature. En effet, de la définition de la distance signée, l'interpénétration de deux corps sera signalée par une distance négative.

L'ensemble  $\mathcal{K}$  est convexe car si  $\mathbf{v}$  et  $\mathbf{w}$  lui appartiennent, alors si  $0 \leq \alpha \leq 1$  :

$$\alpha(\mathbf{v} \cdot \mathbf{n}) \leq \alpha b_0^{\Omega^R}(\mathbf{x}) \text{ et } (1 - \alpha)(\mathbf{w} \cdot \mathbf{n}) \leq (1 - \alpha)b_0^{\Omega^R}(\mathbf{x})$$

de sorte qu'en additionnant :

$$(\alpha\mathbf{v} + (1 - \alpha)\mathbf{w}) \cdot \mathbf{n} \leq b_0^{\Omega^R}(\mathbf{x})$$

De plus, si  $\mathbf{u}$  est la solution recherchée et si  $\mathbf{w} \in \mathcal{K}$  on a sur la zone de contact  $\Gamma_c$  :

$$\begin{aligned} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot (\mathbf{w} - \mathbf{u}) &= \sigma_{nn}(\mathbf{w} \cdot \mathbf{n} - \mathbf{u} \cdot \mathbf{n}) && \text{car } \boldsymbol{\sigma}_T = 0 \\ &= \sigma_{nn}(\mathbf{w} \cdot \mathbf{n} - b_0^{\Omega^R}(\mathbf{x}) + b_0(\mathbf{x}) - \mathbf{u} \cdot \mathbf{n}) \\ &= \sigma_{nn}(\mathbf{w} \cdot \mathbf{n} - b_0^{\Omega^R}(\mathbf{x})) + \sigma_{nn}(b_0^{\Omega^R}(\mathbf{x}) - \mathbf{u} \cdot \mathbf{n}) \\ &= \sigma_{nn}(\mathbf{w} \cdot \mathbf{n} - b_0^{\Omega^R}(\mathbf{x})) && \text{car } (\sigma_{nn}(b_0(\mathbf{x}) - \mathbf{u} \cdot \mathbf{n}) = 0) \\ &\geq 0 && \text{car } \mathbf{w} \in \mathcal{K} \text{ et } \sigma_{nn} \leq 0 \end{aligned}$$

de sorte que l'on peut formuler le problème de contact unilatéral :

Trouver  $\mathbf{u} \in \mathcal{K}$  tel que :

$$\int_{\Omega} (\mathbf{C} : \boldsymbol{\gamma}(\mathbf{u})) : (\boldsymbol{\gamma}(\mathbf{w}) - \boldsymbol{\gamma}(\mathbf{u})) \, dv \geq \int_{\Omega} \mathbf{r} \cdot (\mathbf{w} - \mathbf{u}) \, dv + \int_{\Gamma_N} \mathbf{h} \cdot (\mathbf{w} - \mathbf{u}) \, ds \quad \forall \mathbf{w} \in \mathcal{K} \quad (15.5)$$

Ce type de problème est une inéquation variationnelle et on montre facilement (voir par exemple Youbissi [38]) que ce problème est équivalent à déterminer  $\mathbf{u} \in \mathcal{K}$  solution de :

$$\min_{\mathbf{w} \in \mathcal{K}} \frac{1}{2} a(\mathbf{w}, \mathbf{w}) - L(\mathbf{w}) \quad (15.6)$$

où :

$$a(\mathbf{u}, \mathbf{w}) = \int_{\Omega} (\mathbf{C} : \boldsymbol{\gamma}(\mathbf{u})) : (\boldsymbol{\gamma}(\mathbf{w}) - \boldsymbol{\gamma}(\mathbf{u})) \, dv \text{ et } L(\mathbf{w}) = \int_{\Omega} \mathbf{r} \cdot (\mathbf{w} - \mathbf{u}) \, dv + \int_{\Gamma_N} \mathbf{h} \cdot \mathbf{w} \, ds$$

]

### 15.3 Contact unilatéral en grandes déformations

Nous nous intéressons maintenant au contact unilatéral en grandes déformations, problème beaucoup plus difficile que le précédent. On ne peut plus supposer que le point de contact  $\mathbf{y}(\mathbf{x})$  et la normale  $\mathbf{n}$  ne dépendent pas du déplacement. Il faudra donc les linéariser, ce qui n'est pas une mince tâche, surtout dans le cas du contact déformable-déformable.

#### ♣ Lemme 15.18

On a la propriété suivante :

$$\frac{\partial \mathbf{n}^R}{\partial \mathbf{u}} \boldsymbol{\delta}_u \perp \mathbf{n}^R \quad (15.7)$$

**Démonstration :**

Pour la première partie, puisque  $\| \mathbf{n}^R \|^2 = \mathbf{n}^R \cdot \mathbf{n}^R = 1$ , en dérivant de chaque côté, on a :

$$2 \frac{\partial \mathbf{n}^R}{\partial \mathbf{u}} \delta_u \cdot \mathbf{n}^R = 0$$

d'où le résultat. ♣

**◆ Remarque 15.2**

Pour le problème de deux corps déformables en contact, on a :

$$\frac{\partial b^2}{\partial \mathbf{u}} \cdot \delta_u = \frac{d}{d\lambda} (b^2(\mathbf{x}, \mathbf{u} + \lambda \delta_u - \mathbf{u}_2 - \lambda \delta_u^2))|_{\lambda=0} = \frac{d}{d\lambda} (b_0^2(\mathbf{x}) + (\mathbf{u} + \lambda \delta_u \mathbf{u}_2 - \lambda \delta_u^2) \cdot \mathbf{n}^2)|_{\lambda=0} = \mathbf{n}^2 \cdot (\delta_u - \delta_u^2)$$

◆

**15.3.1 Méthode de pénalisation du contact unilatéral**

La méthode la plus simple, mais certainement pas la meilleure, permettant d'imposer la contrainte de non-pénétration est la méthode de pénalisation qui consiste à définir :

$$b^{R-} = \begin{cases} b^{\Omega^R} & \text{si } b^{\Omega^R} \leq 0 \\ 0 & \text{si } b^{\Omega^R} \geq 0 \end{cases}$$

et à poser :

$$\sigma_{nn} = r b^{R-}$$

tel qu'illustré également à la figure 15.2. Le paramètre de pénalisation  $r$  doit être choisi très grand ( $r \simeq 10^6$ ) pour que la loi de contact unilatéral soit approximativement respectée. C'est là l'un des points faibles de cette approche qui a pour effet, entre autres choses, d'augmenter significativement le conditionnement de la matrice tangente. Nous reviendrons sur ce point plus tard.

Dans la formulation variationnelle 14.24, on verra apparaître à droite un terme de la forme :

$$\int_{\Gamma_C^t} (\sigma_{nn}(\mathbf{w} \cdot \mathbf{n}) + \boldsymbol{\sigma}_t \cdot \mathbf{w}_t) da = \int_{\Gamma_C^t} (\sigma_{nn}(\mathbf{w} \cdot \mathbf{n})) da = \int_{\Gamma_C^t} (r b^{R-}(\mathbf{w} \cdot \mathbf{n})) da$$

puisque, pour le moment, nous supposons nulle la contrainte tangentielle  $\boldsymbol{\sigma}_t$ . Ce terme a pour effet de pénaliser fortement l'interpénétration et n'est présent que dans ce cas et est nul autrement.

Il reste à linéariser ce terme, ce qui découlera directement des résultats de la section 15.1.1. De la relation 15.4, on a la contribution suivante à la matrice tangente :

$$r \int_{\Gamma_C^t} (\delta_u \cdot \mathbf{n}^R)(\mathbf{w} \cdot \mathbf{n}) da$$

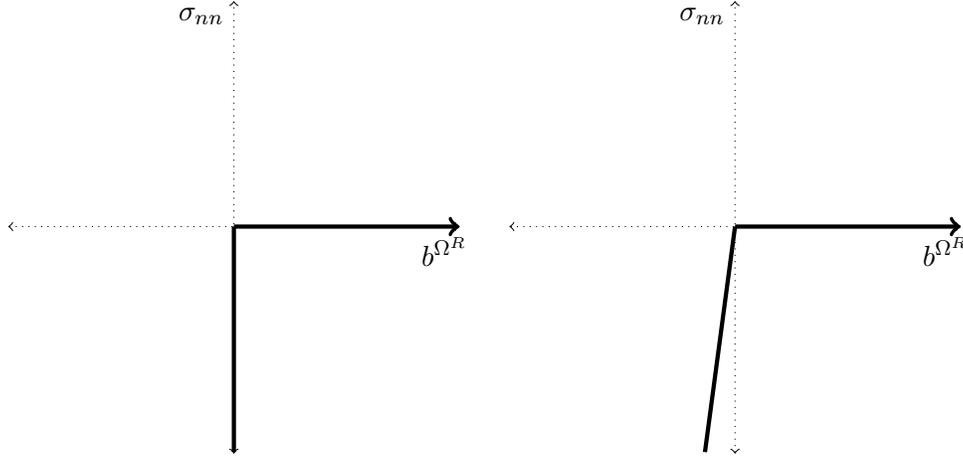


Figure 15.2 – Lois de contact unilatéral originale (à gauche) et pénalisée (à droite)

On remarquera la présence des deux normales  $\mathbf{n}$  et  $\mathbf{n}^R$  qui sont *a priori* différentes. Toutefois, lorsque les deux corps viennent en contact, ces normales sont égales mais de signes opposés  $\mathbf{n} = -\mathbf{n}^R$  et on peut remplacer par :

$$-r \int_{\Gamma_C^t} (\boldsymbol{\delta}_u \cdot \mathbf{n}^R)(\mathbf{w} \cdot \mathbf{n}^R) da = -r \int_{\Gamma_C^0} (\boldsymbol{\delta}_u \cdot \mathbf{n}^R)(\mathbf{w} \cdot \mathbf{n}^R) J_s dA$$

lorsque l'on se ramène à la configuration initiale. Cette contribution est ajoutée à la matrice tangente (avec un changement de signe) et n'est présente que dans la partie de  $\Gamma_C^t$  où il y a interpénétration. La formulation variationnelle devient dans ce cas :

$$\int_{\Omega} (\mathbf{C} : \boldsymbol{\gamma}(\boldsymbol{\delta}_u)) : \boldsymbol{\gamma}(\mathbf{w}) dv - r \int_{\Gamma_C^0} (\boldsymbol{\delta}_u \cdot \mathbf{n}^R)(\mathbf{w} \cdot \mathbf{n}^R) J_s dA = \int_{\Omega} \mathbf{r} \cdot \mathbf{w} dv - \int_{\Omega} (\mathbf{C} : \boldsymbol{\gamma}(\mathbf{u})) : \boldsymbol{\gamma}(\mathbf{w}) dv$$

### ◆ Remarque 15.3

Dans le terme de pénalisation du contact unilatéral, le jacobien surfacique  $J_s$  est souvent omis pour faciliter la mise en oeuvre et pour simplifier la linéarisation (puisque  $J_s$  dépend de  $\mathbf{u}$ ). Les pressions de contact ainsi calculée n'ont donc, *a priori*, pas de signification physique. Par contre, il est relativement facile de récupérer les valeurs physiques par post-traitement. ◆

## 15.3.2 Contact frottant

Le contact frottant est d'une importance capitale en pratique mais constitue tout de même encore un défi sur le plan numérique. La première difficulté provient de la modélisation même

du frottement c.-à-d. de l'élaboration de lois de frottement réalistes. La résolution numérique des équations résultantes est un autre obstacle important à considérer.

Nous nous limiterons dans un premier temps à la loi de frottement de Coulomb classique qui contient déjà les principales difficultés. Cette loi établit un lien direct entre la composante tangentielle du tenseur des contraintes  $\sigma_T$  et la vitesse tangentielle  $\dot{\mathbf{u}}_T$  :

$$\begin{cases} \sigma_T = -\mu|\sigma_{nn}|\frac{\dot{\mathbf{u}}_T}{\|\dot{\mathbf{u}}_T\|} & \text{si } \|\sigma_T\| > \mu|\sigma_{nn}| \\ \dot{\mathbf{u}}_T = 0 & \text{si } \|\sigma_T\| \leq \mu|\sigma_{nn}| \end{cases} \quad (15.8)$$

Le terme  $\frac{\dot{\mathbf{u}}_T}{\|\dot{\mathbf{u}}_T\|}$  est un vecteur unitaire qui donne la direction de la vitesse tangentielle. La force de frottement s'oppose au mouvement, d'où le signe négatif. En dessous d'un certain seuil de contraintes, la vitesse tangentielle est nulle et il y a adhérence entre les deux corps (contact adhérent). Au delà de ce même seuil, la force de frottement est dans la direction inverse de la vitesse tangentielle. La loi de Coulomb est illustrée à la figure 15.3 dans le cas unidimensionnel.

◆ **Remarque 15.4**

L'une des difficultés de mise en oeuvre provient du fait que la loi de Coulomb relie la composante tangentielle du tenseur des contraintes  $\sigma_T$  non pas au déplacement tangentiel mais à la vitesse tangentielle  $\dot{\mathbf{u}}_T$ . Il s'agit-il bien entendu de la vitesse tangentielle relative entre les deux corps en contact c.-à-d. :

$$\dot{\mathbf{u}}_T = (\dot{\mathbf{u}}^1 - \dot{\mathbf{u}}^2)_T$$

Puisque nous utilisons une formulation Lagrangienne totale, on peut remplacer la dérivée temporelle par une différence divisée

$$\dot{\mathbf{u}}_T \simeq \left( \frac{\mathbf{u}^1 - \mathbf{u}_p^1}{\Delta t} - \frac{\mathbf{u}^2 - \mathbf{u}_p^2}{\Delta t} \right)_T$$

où  $\mathbf{u}_p^i, i = 1, 2$  désignent les déplacements au pas de temps précédent. En remplaçant cette dernière expression dans la loi de Coulomb, on remarque que  $\Delta t$  n'a aucune incidence. On peut donc remplacer  $\dot{\mathbf{u}}_T$  par  $(\mathbf{u}^1 - \mathbf{u}_p^1 - (\mathbf{u}^2 - \mathbf{u}_p^2))_T$ . Pour les problèmes dits quasi-statiques, le temps n'est donc pris en compte qu'indirectement, au moyen de ce qu'on appelle des pas de chargements. Dans ce cas,  $\mathbf{u}_p^i$  dénote le déplacement au pas de chargement précédent. Ceci sous-entend une certaine monotonie du chargement, qui n'est cependant pas toujours garantie en pratique.

Dans ce contexte quasi-statique, le déplacement tangentiel, que nous noterons  $\Delta \mathbf{g}_T$ , peut être perçu comme la différence des déplacements entre deux pas de chargement successifs, projetée dans la direction tangentielle à la surface de contact. Les déplacements entre deux pas de chargement de ces deux corps sont notés  $\mathbf{u}^1 - \mathbf{u}_p^1$  et  $\mathbf{u}^2 - \mathbf{u}_p^2$  où l'indice  $p$  désigne les déplacements obtenus au pas de chargement précédent qui sont supposés connus. Le déplacement relatif des deux corps est donc

$$\Delta \mathbf{g} = (\mathbf{u}^1 - \mathbf{u}_p^1) - (\mathbf{u}^2 - \mathbf{u}_p^2)$$

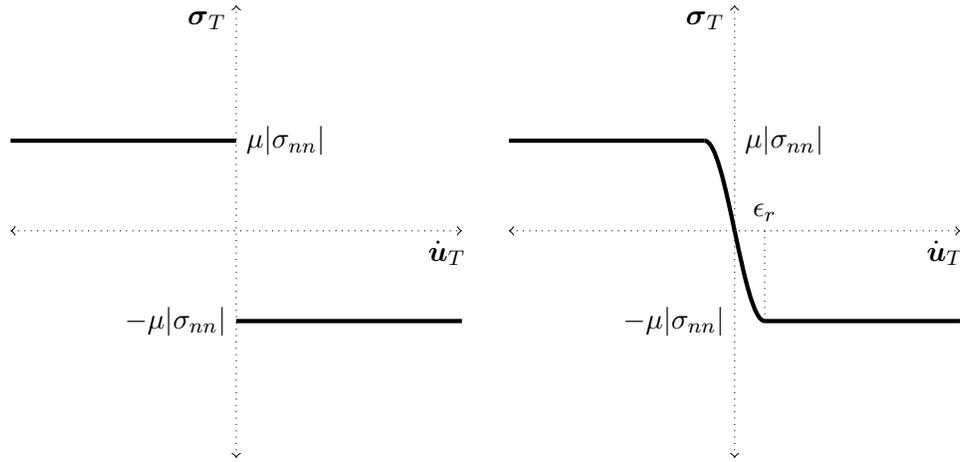


Figure 15.3 – Lois de frottement de Coulomb originale et régularisée

et sa composante tangentielle  $\Delta \mathbf{g}_T$  est donc :

$$\Delta \mathbf{g}_T = (\mathbf{u}^1 - \mathbf{u}_p^1 - (\mathbf{u}^2 - \mathbf{u}_p^2))_T = (\mathbf{u}^1 - \mathbf{u}_p^1) - ((\mathbf{u}^1 - \mathbf{u}_p^1) \cdot \mathbf{n})\mathbf{n} - [(\mathbf{u}^2 - \mathbf{u}_p^2) - ((\mathbf{u}^2 - \mathbf{u}_p^2) \cdot \mathbf{n})\mathbf{n}]$$

où la normale  $\mathbf{n}$  est choisie arbitrairement par  $\mathbf{n} = \mathbf{n}^1 = -\mathbf{n}^2$  et que le choix de cette normale est encore une fois arbitraire.

Dans le cas du frottement sur un corps rigide  $\Omega^R$ , on prendra  $\mathbf{n}^R$  comme normale et on a :

$$\Delta \mathbf{g}_T = (\mathbf{u} - \mathbf{u}_p) - ((\mathbf{u} - \mathbf{u}_p) \cdot \mathbf{n}^R)\mathbf{n}^R - [(\mathbf{u}^R - \mathbf{u}_p^R) - ((\mathbf{u}^R - \mathbf{u}_p^R) \cdot \mathbf{n}^R)\mathbf{n}^R]$$

Le corps rigide peut aussi se déplacer au besoin mais s'il est fixe, on pose tout simplement  $\mathbf{u}^R = \mathbf{u}_p^R = 0$ . C'est donc désormais le déplacement tangentiel relatif des deux corps en mouvement qui entre en jeu. ♦

Tout comme pour le contact unilatéral, on a ici une loi non linéaire et non différentiable qui prévient l'utilisation de méthodes de résolution comme la méthode de Newton. Pour cette raison, on la régularise par l'introduction d'un paramètre  $\epsilon_r$  tel qu'on peut le constater à la figure 15.3. La transition dans l'intervalle  $[0, \epsilon_r]$  peut être simplement linéaire (continue) ou quadratique, ce qui rend la courbe dérivable. C'est ce qui est illustré à la figure 15.3. L'expression de la loi régularisée est alors :

$$\boldsymbol{\sigma}_T = \begin{cases} -\frac{\mu|\sigma_{nn}|}{\epsilon_r} \left(2 - \frac{\|\dot{\mathbf{u}}_T\|}{\epsilon_r}\right) \dot{\mathbf{u}}_T & \text{si } \|\dot{\mathbf{u}}_T\| < \epsilon_r \\ -\mu|\sigma_{nn}| \frac{\dot{\mathbf{u}}_T}{\|\dot{\mathbf{u}}_T\|} & \text{si } \|\dot{\mathbf{u}}_T\| \geq \epsilon_r \end{cases} \quad (15.9)$$

### Évaluation de la vitesse tangentielle

On remarque que lorsque deux corps sont en contact, la composante tangentielle du déplacement ne dépend pas du choix (entre  $\mathbf{n}^1$  et  $\mathbf{n}^2$ ) de la normale. En effet :

$$\mathbf{u}_T^1 = \mathbf{u}^1 - (\mathbf{u}^1 \cdot \mathbf{n}^1)\mathbf{n}^1 = \mathbf{u}^1 - (\mathbf{u}^1 \cdot \mathbf{n}^2)\mathbf{n}^2$$

car  $\mathbf{n}^1 = -\mathbf{n}^2$  et le choix de la normale est arbitraire.

Dans tous les cas, la loi de Coulomb régularisée s'écrit  $\boldsymbol{\sigma}_T = -\mu|\sigma_{nn}|\Psi(\Delta\mathbf{g}_T)$  où :

$$\Psi(\Delta\mathbf{g}_T) = \begin{cases} \frac{1}{\epsilon_r} \left( 2 - \frac{\|\Delta\mathbf{g}_T\|}{\epsilon_r} \right) \Delta\mathbf{g}_T & \text{si } \|\Delta\mathbf{g}_T\| < \epsilon_r \\ \frac{\Delta\mathbf{g}_T}{\|\Delta\mathbf{g}_T\|} & \text{si } \|\Delta\mathbf{g}_T\| \geq \epsilon_r \end{cases} \quad (15.10)$$

### Linéarisation du contact frottant régularisé

La linéarisation du terme de frottement régularisé requiert de dériver par rapport à  $\mathbf{u}$  tous les termes qui en dépendent. Cela inclut  $\Delta\mathbf{g}_T$  mais aussi  $\sigma_{nn}$  lorsque l'on utilise la pénalisation. Enfin, bien que cela soit souvent négligé, il faut inclure la normale  $\mathbf{n}$  qui dépend de  $\mathbf{u}$  comme nous l'avons vu à la relation 14.36. On doit donc linéariser le terme :

$$\int_{\Gamma_C^t} (\boldsymbol{\sigma}_T \cdot \mathbf{w}_T) da = - \int_{\Gamma_C^t} \mu|\sigma_{nn}| \begin{cases} \frac{1}{\epsilon_r} \left( 2 - \frac{\|\Delta\mathbf{g}_T\|}{\epsilon_r} \right) (\Delta\mathbf{g}_T \cdot \mathbf{w}_T) da & \text{si } \|\Delta\mathbf{g}_T\| < \epsilon_r \\ \frac{\Delta\mathbf{g}_T \cdot \mathbf{w}_T}{\|\Delta\mathbf{g}_T\|} da & \text{si } \|\Delta\mathbf{g}_T\| \geq \epsilon_r \end{cases}$$

où  $\boldsymbol{\sigma}_T$  est donné par la loi de Coulomb régularisée 15.10. On doit donc dériver cette expression par rapport à  $\mathbf{u}$ , ce que nous ferons en plusieurs étapes. On a en premier lieu :

$$\begin{aligned} \frac{\partial \Delta\mathbf{g}_T}{\partial \mathbf{u}} \boldsymbol{\delta}_u &= \boldsymbol{\delta}_u^1 - (\boldsymbol{\delta}_u^1 \cdot \mathbf{n}^2)\mathbf{n}^2 - [\boldsymbol{\delta}_u^2 - (\boldsymbol{\delta}_u^2 \cdot \mathbf{n}^1)\mathbf{n}^1] = \boldsymbol{\delta}_{u_T}^1 - \boldsymbol{\delta}_{u_T}^2 = \Delta\mathbf{g}'_T \\ \frac{\partial \|\Delta\mathbf{g}_T\|}{\partial \mathbf{u}} \boldsymbol{\delta}_u &= \frac{\partial (\Delta\mathbf{g}_T \cdot \Delta\mathbf{g}_T)^{1/2}}{\partial \mathbf{u}} \boldsymbol{\delta}_u = \frac{\Delta\mathbf{g}_T \cdot \Delta\mathbf{g}'_T}{\|\Delta\mathbf{g}_T\|} \\ \frac{\partial |\sigma_{nn}|}{\partial \mathbf{u}} \boldsymbol{\delta}_u &= r(\boldsymbol{\delta}_u \cdot \mathbf{n}^R) \end{aligned}$$

La contribution à la matrice tangente se sépare en deux termes.

Si  $\|\Delta\mathbf{g}_T\| < \epsilon_r$ , on a :

$$\begin{aligned}
a(\boldsymbol{\delta}_u, \boldsymbol{w}) &= - \int_{\Gamma_C^t} \frac{\mu |\sigma_{nn}|}{\epsilon_r} \left[ \left( 2 - \frac{\|\Delta \boldsymbol{g}_T\|}{\epsilon_r} \right) (\Delta \boldsymbol{g}'_T \cdot \boldsymbol{w}_T) - \frac{1}{\epsilon_r} \left( \frac{\Delta \boldsymbol{g}_T \cdot \Delta \boldsymbol{g}'_T}{\|\Delta \boldsymbol{g}_T\|} \right) (\Delta \boldsymbol{g}_T \cdot \boldsymbol{w}_T) \right] da \\
&\quad + \int_{\Gamma_C^t} \frac{\mu r}{\epsilon_r} (\boldsymbol{\delta}_u \cdot \boldsymbol{n}^R) \left( 2 - \frac{\|\Delta \boldsymbol{g}_T\|}{\epsilon_r} \right) \Delta \boldsymbol{g}_T \cdot \boldsymbol{w}_T da
\end{aligned}$$

alors que si  $\|\Delta \boldsymbol{g}_T\| \geq \epsilon_r$ , on a :

$$\begin{aligned}
a(\boldsymbol{\delta}_u, \boldsymbol{w}) &= - \int_{\Gamma_C^t} \mu |\sigma_{nn}| \left[ \frac{\Delta \boldsymbol{g}'_T \cdot \boldsymbol{w}_T}{\|\Delta \boldsymbol{g}_T\|} da - \frac{(\Delta \boldsymbol{g}_T \cdot \Delta \boldsymbol{g}'_T)(\Delta \boldsymbol{g}_T \cdot \boldsymbol{w}_T)}{\|\Delta \boldsymbol{g}_T\|^3} da \right] \\
&\quad + \int_{\Gamma_C^t} \mu r (\boldsymbol{\delta}_u \cdot \boldsymbol{n}^R) \frac{(\Delta \boldsymbol{g}_T \cdot \boldsymbol{w}_T)}{\|\Delta \boldsymbol{g}_T\|} da
\end{aligned}$$

## Annexe A

# Rappels sur le théorème de la divergence

Nous rappelons dans cette section quelques résultats fondamentaux sur les fonctions de plusieurs variables réelles et notamment sur le théorème de la divergence qui est à la base de toute formulation variationnelle.

Nous travaillerons généralement dans  $R^3$  ou plus particulièrement dans des ouverts  $\Omega$  de  $R^3$ . Les points de  $\Omega$  sont notés  $\mathbf{x} = (x^1, x^2, x^3)$ , le cas bidimensionnel ne posant évidemment aucune difficulté. En dimension 1, on notera tout simplement  $x$ .

### A.1 Gradient, divergence et laplacien

Rappelons la définition du gradient d'une fonction.

**Définition 1.1**

Soit  $f(\mathbf{x})$  une fonction définie sur un ouvert  $\Omega$ . Le *gradient* de  $f$  en  $\mathbf{x}$  (noté  $\nabla f(\mathbf{x})$ ) est le vecteur pointant dans la direction de croissance maximale de la fonction  $f$  et dont le module est précisément donné par le taux de croissance maximal de  $f$  en  $\mathbf{x}$ .

Cette définition quelque peu étriquée permet de s'affranchir du système de coordonnées. Dans le cas cartésien, on peut montrer que :

$$\nabla f(\mathbf{x}) = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right)$$

**Définition 1.2**

La dérivée directionnelle d'une fonction  $f(\mathbf{x})$  dans la direction du vecteur unitaire  $\mathbf{d}$  ( $\|\mathbf{d}\|_2 = 1$ ) est donnée par :

$$\frac{\partial f}{\partial \mathbf{d}} = \nabla f(\mathbf{x}) \cdot \mathbf{d} = \|\nabla f(\mathbf{x})\|_2 \cos \theta$$

où  $\theta$  est l'angle entre les vecteurs gradient et  $\mathbf{d}$ .

La dérivée directionnelle donne le taux d'accroissement de la fonction  $f(\mathbf{x})$  dans la direction  $\mathbf{d}$ . Lorsque  $\theta = 0$ , les vecteurs gradient et  $\mathbf{d}$  sont dans la même direction et la dérivée directionnelle prend sa valeur maximale ( $\cos \theta = 1$ ), ce qui est bien cohérent avec la définition du gradient..

**Définition 1.3**

Soit un champ de vecteurs  $\mathbf{u}(\mathbf{x})$ . On définit la divergence de ce champ par l'expression :

$$\nabla \cdot \mathbf{u} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3}$$

## A.2 Intégrales curvilignes et surfaciques

### A.2.1 Rappel sur les intégrales curvilignes

Si  $C$  est une courbe de l'espace et  $f(\mathbf{x})$  une fonction scalaire, on pose :

$$\int_C f ds = \int_a^b f(\gamma(t)) \|\gamma'(t)\|_2 dt$$

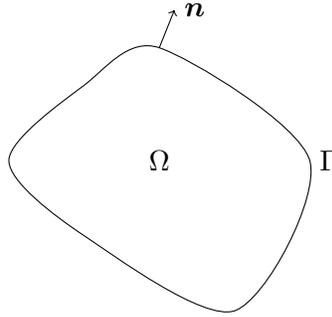
où  $\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$  pour  $a \leq t \leq b$  est une paramétrisation de la courbe  $C$  et  $\|\gamma'(t)\|_2$  désigne la norme euclidienne du vecteur  $\gamma'(t)$  (voir Philippin [28]).

### A.2.2 Rappel sur les intégrales surfaciques

Si  $S$  est une surface de l'espace et  $f(\mathbf{x})$  une fonction scalaire, on pose :

$$\int_S f ds = \int_D f(\boldsymbol{\sigma}(t_1, t_2)) \left\| \frac{\partial \boldsymbol{\sigma}}{\partial t_1} \times \frac{\partial \boldsymbol{\sigma}}{\partial t_2} \right\|_2 dt_1 dt_2$$

où  $\boldsymbol{\sigma}(t_1, t_2) = (\sigma_1(t_1, t_2), \sigma_2(t_1, t_2), \sigma_3(t_1, t_2))$  pour  $(t_1, t_2) \in D$  est une paramétrisation de la surface  $S$  et «  $\times$  » désigne le produit vectoriel. L'intégrale sur la surface  $S$  de la fonction  $f = 1$  donne l'aire

Figure A.1 – Ouvert  $\Omega$  de frontière  $\Gamma$ 

de cette surface. De plus, la normale  $\mathbf{n}$  à la surface est donnée par :

$$\mathbf{n}(t_1, t_2) = \frac{\frac{\partial \boldsymbol{\sigma}}{\partial t_1} \times \frac{\partial \boldsymbol{\sigma}}{\partial t_2}}{\left\| \frac{\partial \boldsymbol{\sigma}}{\partial t_1} \times \frac{\partial \boldsymbol{\sigma}}{\partial t_2} \right\|_2}$$

et une autre forme de ce résultat est que pour une fonction vectorielle  $\mathbf{u}$ , on a :

$$\int_S \mathbf{u} \cdot \mathbf{n} \, ds = \int_D \mathbf{u}(\boldsymbol{\sigma}(t_1, t_2)) \cdot \frac{\partial \boldsymbol{\sigma}}{\partial t_1} \times \frac{\partial \boldsymbol{\sigma}}{\partial t_2} \, dt_1 dt_2$$

### A.3 Théorème de la divergence

Nous sommes en mesure de formuler le théorème de la divergence ou théorème de Gauss-Ostrogradski que nous ne démontrons pas (voir par exemple Swokowski, réf. [36]). Pour un rappel sur les intégrales multiples et les intégrales curvilignes ou surfaciques, on se référera à Philippin [28].

#### ★ Théorème 1.1

Soit  $\Omega$  un ouvert de  $R^n$  de frontière  $\Gamma$  (voir la figure A.1). Soit de plus  $\mathbf{n}$  le vecteur normal unitaire à  $\Gamma$  pointant vers l'extérieur de  $\Omega$ . Si  $\mathbf{u}$  est un champ de vecteurs dont les dérivées partielles premières sont continues sur  $\Omega$ , alors :

$$\int_{\Omega} \nabla \cdot \mathbf{u} \, dv = \int_{\Gamma} \mathbf{u} \cdot \mathbf{n} \, da \tag{A.1}$$

★

#### ◆ Remarque 1.1

Dans le théorème de la divergence, le terme de gauche est une intégrale double ou triple suivant que le problème est en 2 ou 3 dimensions. L'expression  $dv$  désigne donc un élément d'aire ou de volume,

selon le cas. Le terme de droite est une intégrale curviligne si  $\Omega$  est un domaine bidimensionnel ou une intégrale surfacique en 3 dimensions (voir les exercices de fin de chapitre). De même,  $da$  désigne un élément de longueur ou d'aire, selon le cas. Le terme de droite de l'équation A.1 est le *flux* du champ  $\mathbf{u}$  à travers la surface  $\Gamma$ . ♦

Il existe de nombreuses formes sous lesquelles le théorème de la divergence est utile. Nous en explicitons quelques unes.

♠ **Corollaire 1.2**

Si le champ de vecteurs  $\mathbf{u}$  a la forme particulière :

$$\mathbf{u} = (0, 0, \dots, f, 0, \dots, 0)$$

( $f$  apparaissant à la  $i^{\text{ième}}$  composante) où  $f(\mathbf{x})$  est une fonction dont les dérivées partielles premières sont continues, alors la divergence de  $\mathbf{u}$  s'écrit  $\frac{\partial f}{\partial x_i}$  et le théorème de la divergence s'énonce comme suit :

$$\int_{\Omega} \frac{\partial f}{\partial x_i} dv = \int_{\Gamma} f n_i ds \quad (\text{A.2})$$

♠

**Démonstration :**

La preuve vient immédiatement en se servant du théorème de la divergence.

♠ **Corollaire 1.3**

Soit  $h(\mathbf{x})$  une fonction dont les dérivées partielles premières sont continues et si  $\mathbf{u}$  satisfait les hypothèses du théorème de la divergence, alors :

$$\int_{\Omega} h \nabla \cdot \mathbf{u} dv + \int_{\Omega} \nabla h \cdot \mathbf{u} dv = \int_{\Gamma} h \mathbf{u} \cdot \mathbf{n} ds \quad (\text{A.3})$$

En particulier, si le champ de vecteurs  $\mathbf{u}$  a la forme particulière :

$$\mathbf{u} = (0, 0, \dots, f, 0, \dots, 0)$$

( $f$  apparaissant à la  $i^{\text{ième}}$  composante) où  $f(\mathbf{x})$  est une fonction dont les dérivées partielles premières sont continues, alors on a :

$$\int_{\Omega} h \frac{\partial f}{\partial x_i} dv + \int_{\Omega} \frac{\partial h}{\partial x_i} f dv = \int_{\Gamma} h f n_i ds \quad (\text{A.4})$$

♠

**Démonstration :**

Il suffit de démontrer l'identité (voir les exercices de fin de chapitre) :

$$\nabla \cdot (h\mathbf{u}) = h\nabla \cdot \mathbf{u} + \nabla h \cdot \mathbf{u}$$

et d'appliquer directement le théorème de la divergence au champ  $\mathbf{u}_1 = (h\mathbf{u})$ .

La forme la plus répandue de ce théorème est sans doute la suivante.

**♠ Corollaire 1.4**

Si le champ  $\mathbf{u}$  est le produit d'une fonction scalaire  $g$  avec le gradient d'une fonction scalaire  $f(\mathbf{x})$ , c.-à-d.  $\mathbf{u} = g\nabla f$ , alors on a :

$$\int_{\Omega} h\nabla \cdot (g\nabla f) dv + \int_{\Omega} g\nabla h \cdot \nabla f dv = \int_{\Gamma} hg\nabla f \cdot \mathbf{n} ds = \int_{\Gamma} h \left( g \frac{\partial f}{\partial \mathbf{n}} \right) ds \quad (\text{A.5})$$

En particulier, si  $g = 1$ , on a :

$$\int_{\Omega} h\nabla^2 f dv + \int_{\Omega} \nabla h \cdot \nabla f dv = \int_{\Gamma} h \frac{\partial f}{\partial \mathbf{n}} ds \quad (\text{A.6})$$

♠

**Démonstration :**

La démonstration est immédiate en appliquant la relation A.3 et en utilisant la définition du laplacien.

On peut étendre certains des résultats précédents aux tenseurs d'ordre 2. En particulier, on le résultat suivant qui est très utile en élasticité linéaire et en mécanique des fluides.

**♠ Corollaire 1.5**

Si  $\boldsymbol{\sigma}$  est un tenseur symétrique d'ordre 2 et si  $\mathbf{w}$  est un vecteur, alors on a :

$$\int_{\Omega} (\nabla \cdot \boldsymbol{\sigma}) \cdot \mathbf{w} dv + \int_{\Omega} \boldsymbol{\sigma} : \nabla \mathbf{w} dv = \int_{\Gamma} (\boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \mathbf{w} ds \quad (\text{A.7})$$

♠

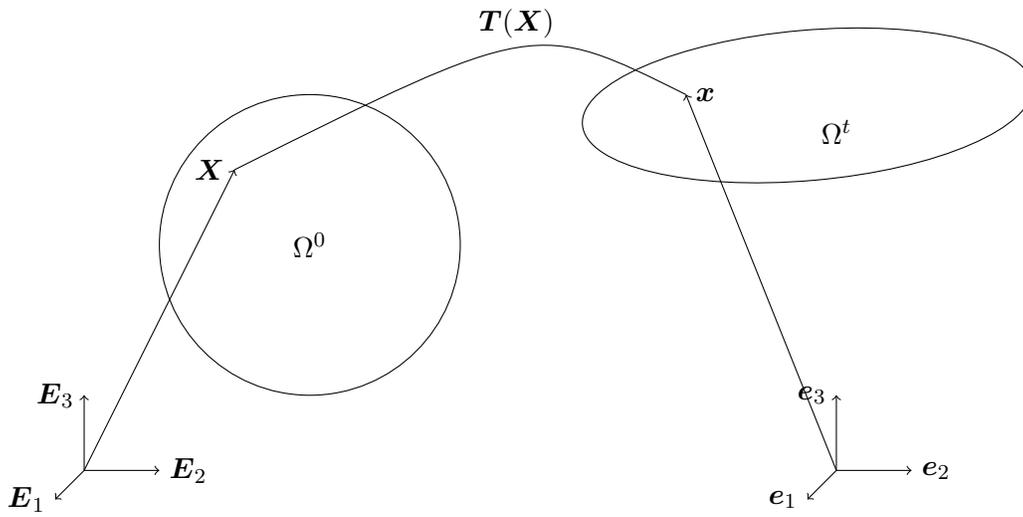


Figure A.2 – Transformation

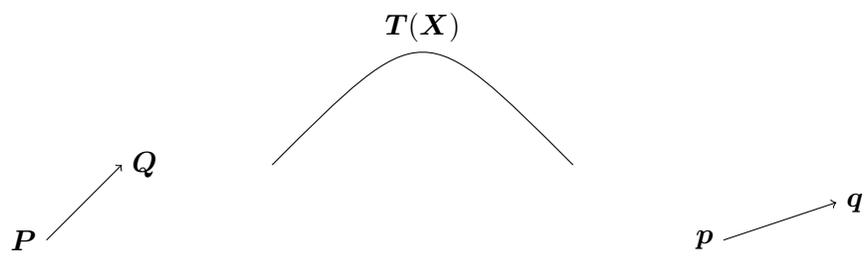


Figure A.3 – Transformation d'un élément de longueur

## A.4 Transformation de Piola

On considère la situation illustrée à la figure A.2. Un domaine  $\Omega^0$  est envoyé sur un autre domaine  $\Omega^t$  par une transformation  $\mathbf{T}$ . On se place dans le cas plus général où des systèmes de coordonnées différents sont utilisés dans  $\Omega^0$  et  $\Omega^t$ . On a donc :

$$\begin{aligned} \mathbf{T} &: \mathbf{X} &\longrightarrow &\mathbf{T}(\mathbf{X}) = \mathbf{x} \\ &: X_i \mathbf{E}_i &\longrightarrow &x_i(\mathbf{X}) \mathbf{e}_i \end{aligned}$$

Soit donc  $\mathbf{P}$  un point de la configuration de référence et posons  $\mathbf{Q} = \mathbf{P} + d\mathbf{X}$ . Ces points seront envoyés respectivement en  $\mathbf{p}$  et  $\mathbf{q} = \mathbf{p} + d\mathbf{x}$  de sorte que :

$$\begin{aligned} d\mathbf{x} &= \mathbf{q} - \mathbf{p} = x_i(\mathbf{X} + d\mathbf{X})\mathbf{e}_i - x_i(\mathbf{X})\mathbf{e}_i = \left( x_i(\mathbf{X}) + \frac{\partial x_i}{\partial X_j} dX_j \right) \mathbf{e}_i - x_i(\mathbf{X})\mathbf{e}_i \\ &= \frac{\partial x_i}{\partial X_j} dX_j \mathbf{e}_i = \frac{\partial x_i}{\partial X_j} (d\mathbf{X} \cdot \mathbf{E}_j) \mathbf{e}_i \\ &= \frac{\partial x_i}{\partial X_j} \mathbf{e}_i \otimes \mathbf{E}_j \cdot d\mathbf{X} = \nabla \mathbf{T} \cdot d\mathbf{X} \end{aligned}$$

Nous avons ainsi introduit le tenseur gradient de déformation  $\nabla \mathbf{T}$  qui gère la déformation d'un élément de longueur. Le tenseur inverse est donné par :

$$\nabla \mathbf{T}^{-1}(\mathbf{x}) = \frac{\partial X_i}{\partial x_j} \mathbf{E}_i \otimes \mathbf{e}_j$$

★ **Théorème 1.2** (Théorème de Piola) :

Soit  $\mathbf{u}$  et  $\boldsymbol{\sigma}$  respectivement un vecteur et un tenseur d'ordre 2 définis dans  $\Omega^t$  c.-à-d.  $\mathbf{u} = u_i \mathbf{e}_i$  et  $\boldsymbol{\sigma} = \sigma_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$ , alors :

$$\int_{\partial \Omega^t} \mathbf{u} \cdot \mathbf{n} \, da = \int_{\partial \Omega^0} \mathbf{U} \cdot \mathbf{N} \, dA \quad \text{et} \quad \int_{\partial \Omega^t} \boldsymbol{\sigma} \cdot \mathbf{n} \, da = \int_{\partial \Omega^0} \mathbf{P} \cdot \mathbf{N} \, dA \quad (\text{A.8})$$

où  $\mathbf{U} = J \nabla \mathbf{T}^{-1} \cdot \mathbf{u}$  et  $\mathbf{P} = J \boldsymbol{\sigma} \nabla \mathbf{T}^{-\top}$ . De plus,  $\nabla \cdot \mathbf{u} = J^{-1} \nabla_X \cdot \mathbf{U}$ .

**Démonstration :**

La première partie du théorème est une conséquence immédiate de la formule de Nanson. Pour la deuxième partie, on considère un élément de volume  $V$  de  $\Omega^0$  qui est envoyé sur un élément  $v$  de  $\Omega^t$ . Du théorème de la divergence, on a d'une part :

$$\int_{\partial v} \mathbf{u} \cdot \mathbf{n} \, da = \int_v \nabla \cdot \mathbf{u} \, dv = \int_V \nabla \cdot \mathbf{u} \, J \, dV$$

et d'autre part :

$$\int_{\partial V} \mathbf{U} \cdot \mathbf{N} \, dA = \int_V \nabla_X \cdot \mathbf{U} \, dV$$

et le résultat suit puisque l'élément de volume  $V$  est quelconque. ★

## A.5 Exercices

1. Démontrer que si  $f(\mathbf{x})$  et  $g(\mathbf{x})$  sont des fonctions scalaires et  $\mathbf{v}(\mathbf{x})$  désigne un champ de vecteurs, on a les relations suivantes :
  - a)  $\nabla \cdot (\nabla f) = \nabla^2 f$
  - b)  $\nabla \cdot (g\mathbf{v}) = \nabla g \cdot \mathbf{v} + g\nabla \cdot \mathbf{v}$
  - c)  $\nabla \cdot (g\nabla f) = \nabla g \cdot \nabla f + g\nabla^2 f$

2. Démontrer que :

a)

$$\int_{\Omega} (f\nabla^2 g - g\nabla^2 f) dv = \int_S \left( f \frac{\partial g}{\partial \mathbf{n}} - g \frac{\partial f}{\partial \mathbf{n}} \right) ds$$

b)

$$\int_{\Omega} \nabla^2 f dv = \int_S \frac{\partial f}{\partial \mathbf{n}} ds$$

### Rappel sur les intégrales curvilignes

Si  $C$  est une courbe de l'espace et  $f(\mathbf{x})$  une fonction scalaire, on pose :

$$\int_C f ds = \int_a^b f(\gamma(t)) \|\gamma'(t)\|_2 dt$$

où  $\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$  pour  $a \leq t \leq b$  est une paramétrisation de la courbe  $C$  et  $\|\gamma'(t)\|_2$  désigne la norme euclidienne du vecteur  $\gamma'(t)$  (voir Philippin [28]).

3. Donner une paramétrisation des courbes suivantes :
  - a) La courbe  $y = x^2$  entre 0 et 1.
  - b) La courbe  $y = f(x)$  entre  $a$  et  $b$ .
  - c) Le cercle de rayon  $R$  et centré au point  $(x_1^c, x_2^c)$ .
  - d) Le segment de droite entre les points  $(1, 2)$  et  $(3, -1)$ .
4. Évaluer les intégrales curvilignes suivantes en donnant au préalable une paramétrisation de la courbe  $C$ .

a)

$$\int_C 1 ds$$

où  $C$  est le cercle de rayon 1 centré en l'origine.

**N.B.** L'intégrale sur la courbe  $C$  de la fonction 1 donne la longueur de cette courbe.

b)

$$\int_C x_1^2 ds$$

où  $C$  est le segment de droite joignant les points  $(1, 1)$  et  $(2, 3)$ .

5. Déterminer, à l'aide de l'intégrale curviligne appropriée, la longueur de l'ellipse :

$$\begin{cases} \gamma_1(t) = a \cos t \\ \gamma_2(t) = b \sin t \end{cases}$$

pout  $t$  entre 0 et  $2\pi$ .

### Rappel sur les intégrales surfaciques

Si  $S$  est une surface de l'espace et  $f(\mathbf{x})$  une fonction scalaire, on pose :

$$\int_S f \, ds = \int_D f(\boldsymbol{\sigma}(t_1, t_2)) \left\| \frac{\partial \boldsymbol{\sigma}}{\partial t_1} \times \frac{\partial \boldsymbol{\sigma}}{\partial t_2} \right\|_2 dt_1 dt_2$$

où  $\boldsymbol{\sigma}(t_1, t_2) = (\sigma_1(t_1, t_2), \sigma_2(t_1, t_2), \sigma_3(t_1, t_2))$  pour  $(t_1, t_2) \in D$  est une paramétrisation de la surface  $S$  et «  $\times$  » désigne le produit vectoriel. L'intégrale sur la surface  $S$  de la fonction  $f = 1$  donne l'aire de cette surface. De plus, la normale  $\mathbf{n}$  à la surface est donnée par :

$$\mathbf{n}(t_1, t_2) = \frac{\frac{\partial \boldsymbol{\sigma}}{\partial t_1} \times \frac{\partial \boldsymbol{\sigma}}{\partial t_2}}{\left\| \frac{\partial \boldsymbol{\sigma}}{\partial t_1} \times \frac{\partial \boldsymbol{\sigma}}{\partial t_2} \right\|_2}$$

et une autre forme de ce résultat est que pour une fonction vectorielle  $\mathbf{u}$ , on a :

$$\int_S \mathbf{u} \cdot \mathbf{n} \, ds = \int_D \mathbf{u}(\boldsymbol{\sigma}(t_1, t_2)) \cdot \frac{\partial \boldsymbol{\sigma}}{\partial t_1} \times \frac{\partial \boldsymbol{\sigma}}{\partial t_2} dt_1 dt_2$$

Voici quelques exemples de paramétrisation de surfaces.

- Toute surface de la forme  $x_3 = f(x_1, x_2)$  pour  $a \leq x_1 \leq b$  et  $c \leq x_2 \leq d$  peut être paramétrisée sous la forme :

$$\boldsymbol{\sigma}(t_1, t_2) = (t_1, t_2, f(t_1, t_2)) \quad a \leq t_1 \leq b, c \leq t_2 \leq d$$

6. Évaluer les intégrales surfaciques suivantes :

a)

$$\int_S (x_1 + x_2 + x_3) \, ds$$

où  $S$  est la surface du plan  $x_3 = x_1 + x_2$  pour  $0 \leq x_2 \leq x_1$  et  $0 \leq x_1 \leq 1$ .

b)

$$\int_S (x_1 + 1) \, ds$$

où  $S$  est la surface définie par  $(\cos t_1, \sin t_1, t_2)$  pour  $0 \leq t_1 \leq 2\pi$  et  $0 \leq t_2 \leq 3$ .

7. Vérifier le théorème de la divergence en calculant séparément l'intégrale de volume et l'intégrale surfacique pour le champ  $\mathbf{v} = (x_1, x_2, x_3)$  et le domaine  $\Omega$  constitué du cube de côté  $a$  :  $[0, a] \times [0, a] \times [0, a]$ .

## Annexe B

# Rappels sur les tenseurs

### B.1 Notions de base

Dans cette annexe, nous ferons un survol très rapide des notions relatives aux tenseurs. L'objectif n'est certainement pas de refaire de manière détaillée toute la théorie tensorielle mais bien de rappeler les principales définitions et les principaux outils : produits dyadiques, contractés, doublement contractés, etc.

#### B.1.1 Les vecteurs

On commence par introduire une base orthonormée  $\mathcal{E} = \{\mathbf{E}_i, i = 1, 2, 3\}$  de  $\mathcal{R}^3$ . Les vecteurs seront notés en caractères gras. Soit donc  $\mathbf{u}$  et  $\mathbf{v}$  des vecteurs de  $\mathcal{R}^3$ . On peut donc les écrire :

$$\mathbf{u} = \sum_{i=1}^3 u_i \mathbf{E}_i = u_i \mathbf{E}_i \text{ et } \mathbf{v} = \sum_{i=1}^3 v_i \mathbf{E}_i = v_i \mathbf{E}_i$$

où nous avons introduit la convention d'Einstein sur les indices répétés.

#### Définition 2.1

Le produit scalaire de deux vecteurs  $\mathbf{u}$  et  $\mathbf{v}$  de  $\mathcal{R}^3$  est noté  $\mathbf{u} \cdot \mathbf{v}$  et est défini par :

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos(\mathbf{u}, \mathbf{v})$$

où  $\cos(\mathbf{u}, \mathbf{v})$  désigne le cosinus de l'angle formé par les deux vecteurs

On remarque immédiatement que si on applique cette définition aux vecteurs de la base orthonormée, on trouve que  $\mathbf{E}_i \cdot \mathbf{E}_j = I_{ij}$  de sorte que :

$$\mathbf{u} \cdot \mathbf{v} = (u_i \mathbf{E}_i) \cdot (v_j \mathbf{E}_j) = u_i v_j I_{ij} = u_i v_i$$

Pour un vecteur  $\mathbf{u}$  quelconque, on peut retrouver ses composantes  $u_i$  dans une base orthonormée donnée simplement en prenant le produit scalaire par  $i^{\text{ième}}$  vecteur de la base. En effet :

$$\mathbf{u} \cdot \mathbf{E}_i = u_j \mathbf{E}_j \cdot \mathbf{E}_i = u_j I_{ij} = u_i \quad (\text{B.1})$$

et ces composantes sont généralement représentées dans une matrice colonne :

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Il ne faut jamais perdre de vue que cette notation est liée à la base  $\mathcal{E}$ . Si on se donne une autre base orthonormée  $\mathcal{E}^* = \{\mathbf{E}_i^*, i = 1, 2, 3\}$ , on peut écrire tout vecteur  $\mathbf{u}$  sous la forme :

$$\mathbf{u} = u_i \mathbf{E}_i \text{ mais aussi } \mathbf{u} = u_i^* \mathbf{E}_i^*$$

On aura ainsi deux matrices colonnes de coefficients pour le même vecteur :

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \text{ et } \begin{bmatrix} u_1^* \\ u_2^* \\ u_3^* \end{bmatrix}$$

chacune liée à une base différente. Pour passer de l'une à l'autre, il faut d'abord exprimer chaque nouveau vecteur de base  $\mathbf{E}_i^*$  dans la base  $\mathcal{E}$  de sorte que  $\mathbf{E}_i^* = A_{ij} \mathbf{E}_j$  (la  $i^{\text{ième}}$  ligne de  $A$  contient les composantes de  $\mathbf{E}_i^*$  dans la base  $\mathcal{E}$ ). En vertu de l'équation B.1, on a tout simplement  $A_{ij} = \mathbf{E}_i^* \cdot \mathbf{E}_j$ . On a ainsi :

$$\mathbf{u} = u_i^* \mathbf{E}_i^* = u_i^* A_{ij} \mathbf{E}_j = u_j \mathbf{E}_j$$

de sorte que  $u_j = A_{ij} u_i^*$  ou encore :

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}^\perp \begin{bmatrix} u_1^* \\ u_2^* \\ u_3^* \end{bmatrix}$$

Inversement, on peut exprimer les vecteurs de la base  $\mathcal{E}$  dans la nouvelle base. En posant  $B_{ij} = \mathbf{E}_i \cdot \mathbf{E}_j^*$ , on a

$$\mathbf{E}_i = B_{ij} \mathbf{E}_j^* = B_{ij} A_{jk} \mathbf{E}_k$$

de sorte  $B_{ij} A_{jk} = I_{ik}$  et que les matrices  $B = A^{-1}$ . Puisque  $A_{ij} = \mathbf{E}_i^* \cdot \mathbf{E}_j$  et  $B_{ij} = \mathbf{E}_i \cdot \mathbf{E}_j^*$ , on remarque immédiatement que  $B_{ij} = A_{ji}$  de sorte  $B = A^\top$ .

### B.1.2 Les tenseurs d'ordre 2

**Définition 2.2**

Un tenseur d'ordre 2  $\mathbf{T}$  est une application linéaire de  $\mathcal{R}^3$  dans  $\mathcal{R}^3$ . Le tenseur  $\mathbf{T}$  associe donc à un vecteur  $\mathbf{u}$  un autre vecteur noté  $\mathbf{T} \cdot \mathbf{u}$ .

Un tenseur est donc une application linéaire et on a :

$$\mathbf{T}(\alpha\mathbf{u} + \beta\mathbf{v}) = \mathbf{T} \cdot (\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha\mathbf{T} \cdot \mathbf{u} + \beta\mathbf{T} \cdot \mathbf{v}$$

Un exemple très simple de tenseur d'ordre 2 est le produit dyadique de deux vecteurs dont nous donnons la définition.

**Définition 2.3**

Le produit dyadique de deux vecteurs  $\mathbf{u}$  et  $\mathbf{v}$  est noté  $\mathbf{u} \otimes \mathbf{v}$  et est défini par :

$$(\mathbf{u} \otimes \mathbf{v}) \cdot \mathbf{w} = (\mathbf{v} \cdot \mathbf{w})\mathbf{u} \quad \forall \mathbf{w}$$

En particulier,  $(\mathbf{E}_i \otimes \mathbf{E}_j) \cdot \mathbf{E}_k = \mathbf{E}_i I_{jk}$ .

Le produit dyadique définit donc une application linéaire de  $\mathcal{R}^3$  dans  $\mathcal{R}^3$  et est donc un tenseur d'ordre 2. On remarque de plus qu'en faisant le produit dyadique de deux des vecteurs de la base orthonormée on obtient :

$$(\mathbf{E}_i \otimes \mathbf{E}_j) \cdot \mathbf{u} = (\mathbf{E}_j \cdot \mathbf{u})\mathbf{E}_i = u_j \mathbf{E}_i$$

En vertu de la linéarité, on a  $\mathbf{T} \cdot \mathbf{u} = \mathbf{T} \cdot (u_j \mathbf{E}_j) = u_j \mathbf{T} \cdot \mathbf{E}_j$  et comme nous l'avons vu, la  $i^{\text{ième}}$  composante de ce vecteur est

$$\mathbf{E}_i \cdot (u_j \mathbf{T} \cdot \mathbf{E}_j) = u_j (\mathbf{E}_i \cdot (\mathbf{T} \cdot \mathbf{E}_j)) = T_{ij} u_j$$

où l'on a posé  $T_{ij} = (\mathbf{E}_i \cdot (\mathbf{T} \cdot \mathbf{E}_j))$ . On a ainsi :

$$\mathbf{T} \cdot \mathbf{u} = T_{ij} u_j \mathbf{E}_i = T_{ij} (\mathbf{E}_i \otimes \mathbf{E}_j) \cdot \mathbf{u}$$

Puisque le vecteur  $\mathbf{u}$  est arbitraire, on a :

$$\mathbf{T} = T_{ij} (\mathbf{E}_i \otimes \mathbf{E}_j) \quad \text{où} \quad T_{ij} = (\mathbf{E}_i \cdot (\mathbf{T} \cdot \mathbf{E}_j)) \quad (\text{B.2})$$

Les coefficients  $T_{ij}$  sont les coefficients du tenseur  $\mathbf{T}$  dans la base orthonormée. En fait, pour définir complètement un tenseur il suffit de connaître son comportement sur les vecteurs de la base orthonormée. Ainsi, en ce qui concerne le produit dyadique de deux vecteurs, on a :

$$(\mathbf{E}_i \cdot ((\mathbf{u} \otimes \mathbf{v}) \cdot \mathbf{E}_j)) = (\mathbf{E}_i \cdot \mathbf{u}) v_j = u_i v_j$$

de sorte que  $\mathbf{u} \otimes \mathbf{v} = u_i v_j \mathbf{E}_i \otimes \mathbf{E}_j$ .

Notons enfin qu'il est possible d'utiliser deux bases orthonormées différentes pour définir un tenseur. En effet, si on introduit une nouvelle base  $\mathbf{E}_i, i = 1, 2, 3$  reliée à la première par un tenseur orthogonal (de rotation)  $\mathbf{A}$  c.-à-d.  $\mathbf{e}_k = \mathbf{A} \cdot \mathbf{E}_k$ . Les composantes de ce tenseur sont donc :

$$A_{ij} = \mathbf{E}_i \cdot (\mathbf{A} \cdot \mathbf{E}_j)$$

on a alors :

$$\mathbf{T} = T_{ij}(\mathbf{e}_i \otimes \mathbf{E}_j) \quad \text{où} \quad T_{ij} = (\mathbf{e}_i \cdot (\mathbf{T} \cdot \mathbf{E}_j)) \quad (\text{B.3})$$

Cette dernière relation sera notamment utile pour les changements de coordonnées.

On peut aussi percevoir les tenseurs d'ordre 2 comme des applications bilinéaires de  $\mathcal{R}^3 \times \mathcal{R}^3$  dans  $\mathcal{R}$  définie par :

$$\mathbf{T}(\mathbf{u}, \mathbf{v}) = (\mathbf{T} \cdot \mathbf{u}) \cdot \mathbf{v} = (T_{ij} u_j \mathbf{e}_i) \cdot \mathbf{v} = T_{ij} u_j v_i \quad \forall \mathbf{u}, \mathbf{v}$$

#### Définition 2.4

Si  $\mathbf{T}$  et  $\mathbf{S}$  désignent des tenseurs d'ordre 2 et  $\mathbf{u}$  et  $\mathbf{v}$  des vecteurs, alors :

1. **Addition** de 2 tenseurs :

$$(\mathbf{T} + \mathbf{S}) \cdot \mathbf{u} = \mathbf{T} \cdot \mathbf{u} + \mathbf{S} \cdot \mathbf{u}$$

2. **Produit** de 2 tenseurs :

$$(\mathbf{T} \cdot \mathbf{S}) \cdot \mathbf{u} = \mathbf{T} \cdot (\mathbf{S} \cdot \mathbf{u})$$

3. **Inverse** d'un tenseur :

$$(\mathbf{T} \cdot \mathbf{T}^{-1}) \cdot \mathbf{u} = \mathbf{T} \cdot (\mathbf{T}^{-1} \cdot \mathbf{u}) = \mathbf{T}^{-1} \cdot (\mathbf{T} \cdot \mathbf{u}) = \mathbf{I} \cdot \mathbf{u} = \mathbf{u}$$

4. La **transposée**  $\mathbf{T}^\top$  d'un tenseur  $\mathbf{T}$  est l'application linéaire notée  $\mathbf{T}^\top$  et vérifiant :

$$(\mathbf{T} \cdot \mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (\mathbf{T}^\top \cdot \mathbf{v}) \quad \forall \mathbf{u}, \mathbf{v}$$

On vérifie facilement que :

$$\mathbf{T}^\top = T_{ji}(\mathbf{e}_i \otimes \mathbf{e}_j)$$

**Définition 2.5**

Le symbole  $\cdot$  désigne le produit contracté (sommation sur l'indice du milieu) de deux tenseurs. Ainsi si  $\mathbf{A}$  et  $\mathbf{B}$  sont des tenseurs du deuxième ordre,  $\mathbf{C}$  un tenseur du quatrième ordre et  $\mathbf{u}$  et  $\mathbf{v}$  des tenseurs du premier ordre (des vecteurs) le produit contracté est défini par :

$$\begin{aligned}(\mathbf{A} \cdot \mathbf{B})_{ij} &= A_{ik} B_{kj} && \text{(un tenseur d'ordre 2)} \\ (\mathbf{C} \cdot \mathbf{B})_{ijkl} &= C_{ijkm} B_{ml} && \text{(un tenseur d'ordre 4)} \\ \mathbf{u} \cdot \mathbf{v} &= u_k v_k && \text{(un scalaire)}\end{aligned}$$

Dans le cas de tenseurs du premier ordre (des vecteurs), on retrouve le produit scalaire habituel. Notons que l'ordre du tenseur résultant du produit contracté est la somme des ordres des tenseurs de départ moins deux.

**Définition 2.6**

Le symbole  $\colon$  désigne le produit doublement contracté (sommation sur les deux indices du milieu) de deux tenseurs. Par exemple, si  $\mathbf{A}$  et  $\mathbf{B}$  sont des tenseurs du second ordre, on a :

$$\mathbf{T} \colon \mathbf{S} = T_{ij} S_{ij}$$

Si  $\mathbf{C}$  est un tenseur du quatrième ordre, alors :

$$\mathbf{C} \colon \mathbf{T} = C_{ijkl} T_{kl}$$

et désigne un tenseur du deuxième ordre. Notons que l'ordre du tenseur résultant du produit doublement contracté est la somme des ordres des tenseurs de départ moins quatre.

**Définition 2.7**

On définit la *trace d'un tenseur* du second ordre par l'expression :

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n A_{ii} \tag{B.4}$$

**★ Théorème 2.1 Propriétés de la trace et du produit contracté**

Si  $\mathbf{A}$ ,  $\mathbf{B}$  et  $\mathbf{C}$  sont des tenseurs du deuxième ordre, alors :

1. Trace du produit :

$$\text{tr}(\mathbf{A} \cdot \mathbf{B}) = \text{tr}(\mathbf{B} \cdot \mathbf{A}) \quad (\text{B.5})$$

2. Trace et produit doublement contracté :

$$\text{tr}(\mathbf{A}) = \mathbf{I} : \mathbf{A} = \mathbf{A} : \mathbf{I} \quad (\text{B.6})$$

3. Orthogonalité :

$$\mathbf{A} : \mathbf{B} = 0 \text{ si } \mathbf{A}^\top = \mathbf{A} \text{ et } \mathbf{B}^\top = -\mathbf{B} \quad (\text{B.7})$$

4.

$$(\mathbf{A} \otimes \mathbf{B}) : \mathbf{C} = \mathbf{A}(\mathbf{B} : \mathbf{C}) \quad (\text{B.8})$$

5. Trace et produit doublement contracté :

$$\mathbf{A} : \mathbf{B} = \text{tr}(\mathbf{A}^\top \cdot \mathbf{B}) = \text{tr}(\mathbf{B}^\top \cdot \mathbf{A}) = \text{tr}(\mathbf{A} \cdot \mathbf{B}^\top) = \text{tr}(\mathbf{B} \cdot \mathbf{A}^\top) \quad (\text{B.9})$$

6.

$$(\mathbf{A} \cdot \mathbf{B}) : \mathbf{C} = \mathbf{B} : (\mathbf{A}^\top \cdot \mathbf{C}) \quad (\text{B.10})$$

★

## B.2 Calcul variationnel avec les tenseurs

Lors du processus de linéarisation des différentes formulations variationnelles des problèmes en grandes déformations, on doit dériver des expressions tensorielles par rapport à  $\mathbf{C}$ ,  $\mathbf{E}$  ou tout simplement par rapport au déplacement  $\mathbf{u}$ . On trouvera dans ce qui suit un certain nombre de résultats utiles à cette fin.

### B.2.1 Résultats généraux

Pour dériver par rapport à un tenseur, on utilise la dérivée de Gâteaux 8.10 introduite au chapitre 8.

#### Définition 2.8

Soit  $f : \mathbf{A} \rightarrow f(\mathbf{A}) \in \mathcal{R}$ . Alors la dérivée de  $f$  par rapport à  $\mathbf{A}$  est un tenseur d'ordre 2 (symétrique si  $\mathbf{A}$  l'est) défini par :

$$\frac{\partial(f(\mathbf{A}))}{\partial \mathbf{A}} : \delta \mathbf{A} = \left. \frac{d}{d\epsilon} f(\mathbf{A} + \epsilon \delta \mathbf{A}) \right|_{\epsilon=0} \quad (\text{B.11})$$

♣ **Lemme 2.19 Règles de base :**

Soit  $f$  une fonction qui associe à un tenseur d'ordre un scalaire et soit  $\mathbf{T}$  une fonction qui associe à un tenseur d'ordre 2 un autre tenseur d'ordre 2. On a alors :

$$\begin{aligned}\frac{\partial(f(\mathbf{A})T(\mathbf{A}))}{\partial\mathbf{A}} &= \mathbf{T}(\mathbf{A}) \otimes \frac{\partial f(\mathbf{A})}{\partial\mathbf{A}} + f(\mathbf{A}) \frac{\partial\mathbf{T}(\mathbf{A})}{\partial\mathbf{A}} \\ \frac{\partial(T_1(\mathbf{A}):T_2(\mathbf{A}))}{\partial\mathbf{A}} &= \mathbf{T}_2(\mathbf{A}) : \frac{\partial T_1(\mathbf{A})}{\partial\mathbf{A}} + \mathbf{T}_1(\mathbf{A}) : \frac{\partial T_2(\mathbf{A})}{\partial\mathbf{A}}\end{aligned}\tag{B.12}$$

♣

♣ **Lemme 2.20**

Si  $\mathbf{B}$  un tenseur d'ordre 2 :

$$\frac{\partial(\text{tr}\mathbf{B})}{\partial\mathbf{B}} = \mathbf{I}\tag{B.13}$$

De plus :

$$\frac{\partial(\text{dét } \mathbf{B})}{\partial\mathbf{B}} = (\text{dét } \mathbf{B}) \mathbf{B}^{-\top}\tag{B.14}$$

et en particulier :

$$\frac{\partial J}{\partial\mathbf{F}} = J \mathbf{F}^{-\top} \quad \text{et} \quad \frac{\partial J}{\partial\mathbf{C}} = \frac{1}{2} J \mathbf{C}^{-1}\tag{B.15}$$

**Démonstration :**

Soit donc  $\mathbf{B}$  un tenseur d'ordre 2. L'opérateur trace étant linéaire, on a :

$$\begin{aligned}\frac{\partial(\text{tr}\mathbf{B})}{\partial\mathbf{B}} : \boldsymbol{\delta}_B &= \left[ \frac{d}{d\epsilon} \text{tr}(\mathbf{B} + \epsilon\boldsymbol{\delta}_B) \right]_{\epsilon=0} \\ &= \text{tr}(\boldsymbol{\delta}_B) = \mathbf{I} : \boldsymbol{\delta}_B\end{aligned}$$

d'où le premier résultat. De plus, par définition du déterminant :

$$\begin{aligned}\frac{\partial(\text{dét } \mathbf{B})}{\partial\mathbf{B}} : \boldsymbol{\delta}_B &= \left[ \frac{d}{d\epsilon} \text{dét } (\mathbf{B} + \epsilon\boldsymbol{\delta}_B) \right]_{\epsilon=0} \\ &= \text{dét } \mathbf{B} \left[ \frac{d}{d\epsilon} \text{dét } (\mathbf{I} + \epsilon\mathbf{B}^{-1} \cdot \boldsymbol{\delta}_B) \right]_{\epsilon=0}\end{aligned}$$

Si on pose  $\mathbf{A} = \mathbf{B}^{-1} \cdot \boldsymbol{\delta}_B$  et que l'on note  $\lambda_i^A$  ses valeurs propres, alors les valeurs propres de  $\epsilon \mathbf{B}^{-1} \cdot \boldsymbol{\delta}_B$  seront  $\epsilon \lambda_i^A$  de sorte que :

$$\det(\epsilon \mathbf{A} - \lambda \mathbf{I}) = (\epsilon \lambda_1^A - \lambda)(\epsilon \lambda_2^A - \lambda)(\epsilon \lambda_3^A - \lambda)$$

et qu'en prenant  $\lambda = -1$ , on a :

$$\det(\epsilon \mathbf{A} + \mathbf{I}) = (\epsilon \lambda_1^A + 1)(\epsilon \lambda_2^A + 1)(\epsilon \lambda_3^A + 1)$$

On montre ensuite facilement que :

$$\frac{d}{d\epsilon} \det(\epsilon \mathbf{A} + \mathbf{I})|_{\epsilon=0} = (\lambda_1^A + \lambda_2^A + \lambda_3^A) = \text{tr}(\mathbf{A}) = \text{tr}(\mathbf{B}^{-1} \cdot \boldsymbol{\delta}_B) = \mathbf{B}^{-\top} : \boldsymbol{\delta}_B$$

en vertu de la relation B.6 et le résultat est donc démontré. La première relation de l'équation B.15 est un cas particulier de ce lemme et pour la deuxième on a :

$$\frac{\partial J}{\partial \mathbf{C}} = \frac{\partial I_3^{1/2}}{\partial \mathbf{C}} = \frac{1}{2} I_3^{-1/2} \frac{\partial I_3}{\partial \mathbf{C}} = \frac{1}{2} I_3^{1/2} \mathbf{C}^{-1}$$

ce qui termine la démonstration. ♣

### ♣ Lemme 2.21

Soit  $\mathbf{B}$  un tenseur d'ordre 2. On a alors :

$$\frac{\partial(\mathbf{B}^{-1})}{\partial \mathbf{B}} : \boldsymbol{\delta}_B = -\mathbf{B}^{-1} \cdot \boldsymbol{\delta}_B \cdot \mathbf{B}^{-1} = -B_{ik}^{-1} B_{lj}^{-1} \delta B_{kl} \quad \text{c.-à-d.} \quad \left( \frac{\partial \mathbf{B}^{-1}}{\partial \mathbf{B}} \right)_{ijkl} = -B_{ik}^{-1} B_{lj}^{-1} \quad (\text{B.16})$$

Comme cas particulier, on a :

$$\frac{\partial(\mathbf{C}^{-1})}{\partial \mathbf{C}} : \boldsymbol{\delta}_C = -\mathbf{C}^{-1} \cdot \boldsymbol{\delta}_C \cdot \mathbf{C}^{-1} = \mathcal{J} : \boldsymbol{\delta}_C$$

où le tenseur du quatrième ordre  $\mathcal{J}$  a pour composantes :

$$\mathcal{J}_{ijkl} = -\frac{1}{2} \left( C_{ik}^{-1} C_{jl}^{-1} + C_{il}^{-1} C_{jk}^{-1} \right)$$

en utilisant la symétrie de  $\mathbf{C}$ . On a aussi symétrisé le tenseur  $\mathcal{C}$  tel que suggéré par Bathe [4].

#### Démonstration :

Puisque  $\mathbf{I} = \mathbf{B} \cdot \mathbf{B}^{-1}$ , en dérivant de chaque côté on a :

$$\begin{aligned} 0 &= \frac{\partial(\mathbf{B} \cdot \mathbf{B}^{-1})}{\partial \mathbf{B}} : \boldsymbol{\delta}_B = \left( \frac{\partial \mathbf{B}}{\partial \mathbf{B}} : \boldsymbol{\delta}_B \right) \cdot \mathbf{B}^{-1} + \mathbf{B} \cdot \left( \frac{\partial \mathbf{B}^{-1}}{\partial \mathbf{B}} : \boldsymbol{\delta}_B \right) \\ &= \boldsymbol{\delta}_B \cdot \mathbf{B}^{-1} + \mathbf{B} \cdot \left( \frac{\partial \mathbf{B}^{-1}}{\partial \mathbf{B}} : \boldsymbol{\delta}_B \right) \end{aligned}$$

ce qui entraîne que :

$$\mathbf{B} \cdot \left( \frac{\partial \mathbf{B}^{-1}}{\partial \mathbf{B}} \cdot \boldsymbol{\delta}_B \right) = -\boldsymbol{\delta}_B \cdot \mathbf{B}^{-1}$$

et le résultat suit par une multiplication à gauche par  $\mathbf{B}^{-1}$  de chaque côté. De plus :

$$-\mathbf{B}^{-1} \cdot \boldsymbol{\delta}_B \cdot \mathbf{B}^{-1} = -B_{ik}^{-1} \delta_{Bkl} B_{lj}^{-1} = -B_{ik}^{-1} B_{lj}^{-1} \delta_{Bkl} = \mathcal{J} : \boldsymbol{\delta}_B$$

quel que soit  $\boldsymbol{\delta}_B$ . ♣

### ♣ Lemme 2.22

On a également :

$$\left( \frac{\partial \mathbf{B}^{-\top}}{\partial \mathbf{B}} \right) = -\mathbf{B}^{-\top} \cdot \boldsymbol{\delta}_B^{\top} \cdot \mathbf{B}^{-\top} = \mathcal{K} : \boldsymbol{\delta}_B \quad \text{avec} \quad \mathcal{K}_{ijkl} = \frac{\partial B_{ij}^{-\top}}{\partial B_{kl}} = -B_{il}^{-\top} B_{kj}^{-\top} \quad (\text{B.17})$$

**Démonstration :**

$$\frac{\partial \mathbf{B}^{-\top}}{\partial \mathbf{B}} : \boldsymbol{\delta}_B = \frac{\partial \mathbf{B}^{-\top}}{\partial \mathbf{B}^{\top}} \cdot \left( \frac{\partial \mathbf{B}^{\top}}{\partial \mathbf{B}} : \boldsymbol{\delta}_B \right) = -\mathbf{B}^{-\top} \cdot \left( \frac{\partial \mathbf{B}^{\top}}{\partial \mathbf{B}} : \boldsymbol{\delta}_B \right) \cdot \mathbf{B}^{-\top} = -\mathbf{B}^{-\top} \cdot \boldsymbol{\delta}_B^{\top} \cdot \mathbf{B}^{-\top} = \mathcal{K} : \boldsymbol{\delta}_B$$

ce qui définit le tenseur du quatrième ordre  $\mathcal{K}_{ijkl} = -B_{il}^{-\top} B_{kj}^{-\top}$ . ♣

## B.2.2 Dérivées des invariants

### ♣ Lemme 2.23

Les dérivées premières des trois invariants de  $\mathbf{C}$  s'écrivent :

$$\begin{aligned} \frac{\partial I_1}{\partial \mathbf{C}} &= \mathbf{I} \\ \frac{\partial I_2}{\partial \mathbf{C}} &= I_1 \mathbf{I} - \mathbf{C} \\ \frac{\partial I_3}{\partial \mathbf{C}} &= I_3 \mathbf{C}^{-\top} = I_3 \mathbf{C}^{-1} = \text{cof } \mathbf{C} \end{aligned} \quad (\text{B.18})$$

**Démonstration pour  $I_1$  :**

$$\begin{aligned} \frac{\partial I_1}{\partial \mathbf{C}} : \delta \mathbf{C} &= \frac{d}{d\epsilon} (\text{tr}(\mathbf{C} + \epsilon \delta \mathbf{C})) \Big|_{\epsilon=0} \\ &= \frac{d}{d\epsilon} (\text{tr}(\mathbf{C}) + \epsilon \text{tr}(\delta \mathbf{C})) \Big|_{\epsilon=0} = \text{tr}(\delta \mathbf{C}) \\ &= \mathbf{I} : \delta \mathbf{C} \end{aligned}$$

d'où le résultat puisque  $\delta \mathbf{C}$  est quelconque.

**Démonstration pour  $I_2$  :**

$$\begin{aligned} \frac{\partial I_2}{\partial \mathbf{C}} : \delta \mathbf{C} &= \frac{1}{2} \left[ 2I_1 \frac{\partial I_1}{\partial \mathbf{C}} : \delta \mathbf{C} - \frac{\partial (\mathbf{C} : \mathbf{C})}{\partial \mathbf{C}} : \delta \mathbf{C} \right] \\ &= I_1 \mathbf{I} : \delta \mathbf{C} - \frac{1}{2} \left[ \frac{d}{d\epsilon} ((\mathbf{C} + \epsilon \delta \mathbf{C}) : (\mathbf{C} + \epsilon \delta \mathbf{C})) \right]_{\epsilon=0} \\ &= I_1 \mathbf{I} : \delta \mathbf{C} - \frac{1}{2} \left[ \frac{d}{d\epsilon} (\mathbf{C} : \mathbf{C} + \epsilon \mathbf{C} : \delta \mathbf{C} + \epsilon \delta \mathbf{C} : \mathbf{C} + \epsilon^2 \delta \mathbf{C} : \delta \mathbf{C}) \right]_{\epsilon=0} \\ &= I_1 \mathbf{I} : \delta \mathbf{C} - \mathbf{C} : \delta \mathbf{C} = (I_1 \mathbf{I} - \mathbf{C}) : \delta \mathbf{C} \end{aligned}$$

et le résultat suit immédiatement.

**Démonstration pour  $I_3$  :**

C'est un cas particulier de la relation [B.14](#). ♣

### ♣ Lemme 2.24

Les dérivées secondes des trois invariants de  $\mathbf{C}$  s'écrivent :

$$\begin{aligned} \frac{\partial^2 I_1}{\partial \mathbf{C} \partial \mathbf{C}} &= \mathbf{0} \\ \frac{\partial^2 I_2}{\partial \mathbf{C} \partial \mathbf{C}} &= \mathbf{I} \otimes \mathbf{I} - \mathcal{I} \\ \frac{\partial^2 I_3}{\partial \mathbf{C} \partial \mathbf{C}} &= I_3 (\mathbf{C}^{-1} \otimes \mathbf{C}^{-1} + \mathcal{J}) \end{aligned} \tag{B.19}$$

où  $\mathcal{I}$  est le tenseur identité du quatrième ordre défini par :

$$\mathcal{I} : \mathbf{B} = \mathbf{B} \quad \forall \mathbf{B} \quad \text{c.-à-d.} \quad \mathcal{I}_{ijkl} = \frac{1}{2} (I_{ik} I_{jl} + I_{il} I_{jk})$$

**Démonstration :**

La démonstration est une conséquence immédiate des lemmes précédents. ♣

♣ **Lemme 2.25**

Dérivées premières des  $J_i$  :

$$\begin{aligned}\frac{\partial J_1}{\partial \mathbf{C}} &= \frac{\partial I_1}{\partial \mathbf{C}} I_3^{-1/3} - \frac{1}{3} I_1 I_3^{-4/3} \frac{\partial I_3}{\partial \mathbf{C}} = I_3^{-1/3} \mathbf{I} - \frac{1}{3} I_1 I_3^{-1/3} \mathbf{C}^{-1} \\ \frac{\partial J_2}{\partial \mathbf{C}} &= \frac{\partial I_2}{\partial \mathbf{C}} I_3^{-2/3} - \frac{2}{3} I_2 I_3^{-5/3} \frac{\partial I_3}{\partial \mathbf{C}} = (I_1 \mathbf{I} - \mathbf{C}) I_3^{-2/3} - \frac{2}{3} I_2 I_3^{-2/3} \mathbf{C}^{-1}\end{aligned}\tag{B.20}$$

**Démonstration :** Le résultat est immédiat. ♣

### B.2.3 Application aux grandes déformations

♣ **Lemme 2.26**

$$\frac{\partial \mathbf{F}}{\partial \mathbf{u}} \cdot \mathbf{w} = \nabla_X \mathbf{w} \quad \text{et} \quad \frac{\partial \mathbf{F}^\top}{\partial \mathbf{u}} \cdot \mathbf{w} = \nabla_X^\top \mathbf{w}\tag{B.21}$$

**Démonstration :**

On a immédiatement que :

$$\frac{\partial \mathbf{F}(\mathbf{u})}{\partial \mathbf{u}} \cdot \mathbf{w} = \frac{d}{d\epsilon} [\mathbf{F}(\mathbf{u} + \epsilon \mathbf{w})]_{\epsilon=0} = \frac{d}{d\epsilon} [\mathbf{F}(\mathbf{u}) + \epsilon \nabla_X \mathbf{w}]_{\epsilon=0} = \nabla_X \mathbf{w}$$

La deuxième relation se démontre de manière identique. ♣

On tire des résultats précédents une identité qui nous servira à établir le théorème de transport de Reynolds.

♣ **Lemme 2.27**

On a :

$$\frac{\partial J}{\partial t}(\mathbf{X}, t) = J \nabla \cdot \mathbf{v}\tag{B.22}$$

**Démonstration :** En utilisant le lemme précédent et la relation B.15, on a immédiatement que :

$$\left. \frac{\partial J}{\partial t}(\mathbf{X}, t) \right|_{\mathbf{X}} = \frac{\partial J}{\partial \mathbf{F}} : \left( \left. \frac{\partial \mathbf{F}}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{u}}{\partial t} \right|_{\mathbf{X}} \right) = \frac{\partial J}{\partial \mathbf{F}} : \left( \left. \frac{\partial \mathbf{F}}{\partial \mathbf{u}} \cdot \mathbf{v} \right|_{\mathbf{X}} \right) = J \mathbf{F}^{-\top} : \nabla_X \mathbf{v}$$

où  $\mathbf{v}$  est le champ de vitesse. Par la suite :

$$\begin{aligned} \left. \frac{\partial J}{\partial t}(\mathbf{X}, t) \right|_{\mathbf{X}} &= J \mathbf{F}^{-\top} : \nabla_X \mathbf{v} \\ &= J \operatorname{tr} (\mathbf{F}^{-1} \cdot \nabla_X \mathbf{v}) = J \operatorname{tr} (\nabla_X \mathbf{v} \cdot \mathbf{F}^{-1}) \\ &= J \operatorname{tr} (\nabla \mathbf{v}) = J (\nabla \cdot \mathbf{v}) \end{aligned}$$

en utilisant B.9 et on a le résultat. ♣

### ♣ Lemme 2.28

Le tenseur  $\mathbf{E}$  dépend de manière non linéaire du déplacement  $\mathbf{u}$ . On peut donc le linéariser et on a :

$$\frac{\partial \mathbf{E}(\mathbf{u})}{\partial \mathbf{u}} \cdot \mathbf{w} = \frac{1}{2} \left( \mathbf{F}^\top \cdot \nabla_X \mathbf{w} + \nabla_X^\top \mathbf{w} \cdot \mathbf{F} \right) = \mathbf{F}^\top \gamma(\mathbf{w}) \mathbf{F} \quad (\text{B.23})$$

**Démonstration :**

On a par définition :

$$\begin{aligned} \frac{\partial \mathbf{E}}{\partial \mathbf{u}} \cdot \mathbf{w} &= \frac{1}{2} \frac{\partial \mathbf{C}}{\partial \mathbf{u}} \cdot \mathbf{w} = \frac{1}{2} \left( \left( \frac{\partial \mathbf{F}^\top}{\partial \mathbf{u}} \cdot \mathbf{w} \right) \cdot \mathbf{F} + \mathbf{F}^\top \cdot \left( \frac{\partial \mathbf{F}^\top}{\partial \mathbf{u}} \cdot \mathbf{w} \right) \right) \\ &\quad \frac{1}{2} \left( \nabla_X^\top \mathbf{w} \cdot \mathbf{F} + \mathbf{F}^\top \cdot \nabla_X \mathbf{w} \right) \end{aligned}$$

en vertu du lemme précédent. La dernière égalité de B.23 découle de la relation 14.5. ♣

### ◆ Remarque 2.1

On montre facilement que :

$$\frac{\partial \gamma(\mathbf{u})}{\partial \mathbf{u}} \cdot \mathbf{w} = \gamma(\mathbf{w})$$

On utilise de plus très souvent la notation :

$$\delta \mathbf{E} = \frac{\partial \mathbf{E}}{\partial \mathbf{u}} \cdot \mathbf{w} \quad \text{et} \quad \delta \mathbf{e} = \frac{\partial \gamma(\mathbf{u})}{\partial \mathbf{u}} \cdot \mathbf{w}$$

ce qui permet d'écrire :

$$\delta \mathbf{E} = \mathbf{F}^\top \cdot \delta \mathbf{e} \cdot \mathbf{F}$$

Nous avons choisi d'éviter cette notation qui, bien que très compacte, escamote quelque peu les dépendances des différents termes par rapport à  $\mathbf{u}$  et  $\mathbf{w}$ . ◆

♣ **Lemme 2.29**

$$\frac{\partial J}{\partial \mathbf{u}} \cdot \mathbf{w} = \frac{\partial J}{\partial \mathbf{F}} : \left( \frac{\partial \mathbf{F}}{\partial \mathbf{u}} \cdot \mathbf{w} \right) = J \mathbf{F}^{-\top} : \nabla_X \mathbf{w} = J \nabla \cdot \mathbf{w} \quad (\text{B.24})$$

**Démonstration :**

La démonstration découle des relations [B.15](#), [B.21](#) et [14.4](#) et du fait que :

$$\mathbf{F}^{-\top} : \nabla_X \mathbf{w} = \text{tr}(\mathbf{F}^{-1} \cdot \nabla_X \mathbf{w}) = \text{tr}(\nabla \mathbf{w}) = \nabla \cdot \mathbf{w}$$

♣

### B.3 Exercices

1. Vérifier les identités suivantes pour  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  et  $\mathbf{D}$  des tenseurs d'ordre 2 et  $\mathcal{T}$  un tenseur d'ordre 4.
  - a)  $\mathbf{A}:(\mathbf{B}\cdot\mathbf{C}) = (\mathbf{A}\cdot\mathbf{C}^\top):\mathbf{B} = (\mathbf{B}^\top\cdot\mathbf{A}):\mathbf{C}$
  - b)  $\mathcal{T}:(\mathbf{B}\cdot\mathbf{C}) = (\mathcal{T}\cdot\mathbf{C}^\top):\mathbf{B}$  mais  $\neq (\mathbf{B}^\top\cdot\mathcal{T}):\mathbf{C}$
  - c)  $\mathbf{A}:\mathbf{B} = \mathbf{B}:\mathbf{A}$
  - d)  $(\mathcal{T}:\mathbf{B}):\mathbf{C} = \mathbf{C}:(\mathcal{T}:\mathbf{B}) = (\mathbf{C}:\mathcal{T}):\mathbf{B}$
  - e)  $(\mathcal{T}:(\mathbf{A}\cdot\mathbf{B})):(\mathbf{C}\cdot\mathbf{D}) = ((\mathbf{C}^\top\cdot\mathcal{T}\cdot\mathbf{B}^\top):\mathbf{A}):\mathbf{D}$
  - f) Si  $\mathcal{T}$  est symétrique, alors  $(\mathcal{T}:\mathbf{A}):\mathbf{B} = (\mathcal{T}:\mathbf{B}):\mathbf{A}$
  - g) Si  $\mathcal{T}$  est symétrique, alors  $(\mathcal{T}:(\mathbf{A}\cdot\mathbf{B})):(\mathbf{C}\cdot\mathbf{D}) = ((\mathbf{A}^\top\cdot\mathcal{T}\cdot\mathbf{D}^\top):\mathbf{C}):\mathbf{B}$
2. Démontrer les identités 14.41 et 14.42 en vous servant des identités du numéro précédent.
3. Soit  $\mathbf{A}$  et  $\mathbf{B}$  des tenseurs d'ordre 2 et supposons que  $\mathbf{A}$  est symétrique. Montrer que :

$$\mathbf{A}:\left(\frac{\mathbf{B} + \mathbf{B}^\top}{2}\right) = \mathbf{A}:\mathbf{B}$$

4. Soit  $\mathcal{T}$  un tenseur symétrique d'ordre 4 et  $\mathbf{A}$  un tenseur d'ordre 2. Montrer que :

$$\mathcal{T}:\left(\frac{\mathbf{A} + \mathbf{A}^\top}{2}\right) = \mathcal{T}:\mathbf{A}$$

5. Soit  $\mathcal{T}$  un tenseur symétrique d'ordre 4 et  $\mathbf{A}$  et  $\mathbf{B}$  deux tenseurs d'ordre 2. Montrer que :

$$\left(\mathcal{T}:\left(\frac{\mathbf{A} + \mathbf{A}^\top}{2}\right)\right):\left(\frac{\mathbf{B} + \mathbf{B}^\top}{2}\right) = (\mathcal{T}:\mathbf{A}):\mathbf{B}$$

6. Montrer que :

$$\frac{\partial(\mathbf{B}:\mathbf{B})}{\partial\mathbf{B}} = 2\mathbf{B}$$

7. Montrer que le tenseur du quatrième ordre  $\mathcal{I}$  ayant pour composantes  $\mathcal{I}_{ijkl} = I_{ik}I_{jl}$  vérifie  $\mathcal{I}:\mathbf{B} = \mathbf{B}$  pour tout tenseur  $\mathbf{B}$  d'ordre 2. Montrer de plus que  $\mathcal{I}\cdot\mathcal{B} = \mathcal{B}$  pour tout tenseur  $\mathcal{B}$  d'ordre 4.

## Annexe C

# Interpolation de Lagrange

Le but de ce chapitre est de rappeler les techniques classiques d'interpolation et en particulier la construction des fonctions d'interpolation de Lagrange. Dans le contexte de la méthode des éléments finis, nous nous limiterons à déterminer les fonctions d'interpolation seulement sur les éléments de référence. Nous procéderons de manière identique en dimension 1, 2 ou 3, même s'il existe des façons plus simples d'arriver au même résultat.

Deux espaces de polynômes jouent un rôle important en éléments finis. Dans un premier temps, nous noterons  $P_k$  l'espace des polynômes de degré  $k$ . On vérifie facilement que la dimension de l'espace  $P_k$  est :

$$\dim(P_k) = \frac{(n+k)!}{k!n!}$$

où  $n$  est la dimension d'espace ( $n = 1, 2$  ou  $3$ ). L'espace  $P_k$  est généralement utilisé sur les éléments triangulaires ou tétraédriques.

Nous travaillerons également avec l'espace  $Q_k$  des polynômes de degré  $k$  en chacune des variables d'espace. La dimension de cet espace est tout simplement :

$$\dim(Q_k) = (k+1)^n$$

En dimension 1, ces 2 espaces coïncident mais ils diffèrent en dimension 2 ou 3. L'espace  $Q_k$  est particulièrement adapté aux éléments quadrangulaires ou hexaédraux comme nous le verrons un peu plus loin dans ce chapitre. Bien sûr, il sera possible d'utiliser des variantes incomplètes des espaces  $P_k$  ou  $Q_k$  selon les besoins.

Quelque soit l'espace de polynômes utilisé, nous noterons  $N$  sa dimension. On se donne maintenant  $N$  points  $\xi_1, \xi_2, \dots, \xi_N$  sur l'élément de référence. Les  $N$  fonctions de Lagrange  $L_i(\xi)$  sont des polynômes de  $P_k$  ou  $Q_k$  vérifiant :

$$\begin{cases} L_i(\xi_i) = 1 & \forall i \\ L_i(\xi_j) = 0 & \forall j \neq i \end{cases} \quad (\text{C.1})$$

Pour construire ces fonctions, il suffit de choisir une base généralement constituée de monômes de degré  $k$ . Nous noterons cette base de polynômes  $p_1(\xi), p_2(\xi), \dots, p_N(\xi)$ . Pour construire les

fonctions C.1, il suffit alors de poser :

$$L_i(\boldsymbol{\xi}) = \sum_{j=1}^N a_j p_j(\boldsymbol{\xi})$$

et d'imposer les contraintes de l'équation C.1 pour construire un système linéaire de dimension  $N$  sur  $N$  dont les inconnues sont les coefficients  $a_i$ . On obtient ainsi le système linéaire :

$$\begin{bmatrix} p_1(\boldsymbol{\xi}_1) & p_2(\boldsymbol{\xi}_1) & \cdots & p_N(\boldsymbol{\xi}_1) \\ p_1(\boldsymbol{\xi}_2) & p_2(\boldsymbol{\xi}_2) & \cdots & p_N(\boldsymbol{\xi}_2) \\ \vdots & \ddots & & \vdots \\ p_1(\boldsymbol{\xi}_i) & p_2(\boldsymbol{\xi}_i) & \cdots & p_N(\boldsymbol{\xi}_i) \\ \vdots & & \ddots & \vdots \\ p_1(\boldsymbol{\xi}_N) & p_2(\boldsymbol{\xi}_N) & \cdots & p_N(\boldsymbol{\xi}_N) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_i \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad (\text{C.2})$$

Le terme de droite est tout simplement le vecteur de base  $\mathbf{e}_i$  dont toutes les composantes sont nulles, sauf la  $i^{\text{ième}}$ . La résolution est immédiate puisque la matrice est généralement inversible. Notons de plus que la matrice est la même pour toutes les fonctions de Lagrange. Seul le terme de droite change.

On répète ce processus pour chaque fonction de Lagrange. Une fois ce travail accompli, le polynôme :

$$u(\boldsymbol{\xi}) = \sum_{j=1}^N u_j L_j(\boldsymbol{\xi}) \quad (\text{C.3})$$

est l'unique polynôme de degré  $k$  dont la valeur en  $\boldsymbol{\xi} = \boldsymbol{\xi}_i$  est  $u_i$ , ce qui constitue un polynôme d'interpolation.

## C.1 Interpolation en dimension 1

En dimension 1, l'élément de référence est l'intervalle  $] -1, 1[$ . La dimension de  $P_k$  est tout simplement  $N = k+1$  et on choisit une base naturelle de cet espace de polynômes soit  $1, \xi, \xi^2, \dots, \xi^k$ .

### Polynômes de degré 1

On choisit naturellement comme points d'interpolation les coordonnées  $\xi_1 = -1$  et  $\xi_2 = 1$  de l'élément de référence. La base de monômes est constituée des polynômes 1 et  $\xi$ . Le système C.2 devient pour la fonction  $L_1(\xi)$  :

$$\begin{bmatrix} 1 & -1 \\ 1 & +1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

tandis que pour la fonction  $L_2(\xi)$ , le système est :

$$\begin{bmatrix} 1 & -1 \\ 1 & +1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Polynômes de Lagrange $P_1$ (1D)		
$i$	$L_i(\xi, \eta)$	$\frac{dL_i}{d\xi}$
1	$\frac{1 - \xi}{2}$	$-\frac{1}{2}$
2	$\frac{1 + \xi}{2}$	$\frac{1}{2}$

TABLE C.1 – Polynômes de Lagrange  $P_1$  sur l'intervalle de référence

On obtient ainsi les solutions :

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} \text{ et } \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

Ces deux fonctions sont illustrées à la figure 5.4 s'expriment à l'aide de la relation C.3 sous la forme :

$$L_1(\xi) = \frac{1 - \xi}{2} \text{ et } L_2(\xi) = \frac{1 + \xi}{2}$$

Pour compléter le développement, le tableau C.1 donne l'expression des fonctions de Lagrange de même que les dérivées qui sont utiles pour l'évaluation des systèmes élémentaires comme celui de la relation 5.5.

**Polynômes de degré 2**

On choisit naturellement comme points d'interpolation les coordonnées  $\xi_1 = -1$ ,  $\xi_2 = 1$  et  $\xi_3 = 0$  de l'élément de référence (nous avons numéroté les extrémités ou noeuds géométriques de l'élément en premier). La base de monômes est constituée des polynômes 1,  $\xi$  et  $\xi^2$ . Nous ne ferons explicitement le calcul que pour la première fonction de Lagrange. Le système C.2 s'écrit alors :

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & +1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

dont la solution est :

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix}$$

Polynômes de Lagrange $P_2$ (1D)		
$i$	$L_i(\xi, \eta)$	$\frac{dL_i}{d\xi}$
1	$\frac{\xi(\xi - 1)}{2}$	$\xi - 1/2$
2	$\frac{\xi(\xi + 1)}{2}$	$\xi + 1/2$
3	$1 - \xi^2$	$-2\xi$

TABLE C.2 – Polynômes de Lagrange  $P_2$  sur l'intervalle de référence

et en vertu de l'équation C.3, on a donc la fonction :

$$L_1(\xi) = 0 - \frac{\xi}{2} + \frac{\xi^2}{2} = \frac{\xi(\xi - 1)}{2}$$

De même, on obtient les autres fonctions de Lagrange listées au tableau C.2.

◆ **Remarque 3.1**

On peut procéder plus simplement et construire les polynômes de Lagrange directement. Pour obtenir par exemple  $L_1(\xi)$  en se sert directement de l'équation C.1. En effet, cette fonction doit s'annuler en  $\xi = 0$  et en  $\xi = 1$ . Il suffit donc d'introduire des facteurs  $\xi$  et  $(\xi - 1)$ . De plus, elle doit prendre la valeur 1 en  $\xi = -1$ . Or la fonction  $\xi(\xi - 1)$  vaut 2 à cette endroit. Il suffit donc de diviser par 2 et on obtient :

$$L_1(\xi) = \frac{1}{2}\xi(\xi - 1)$$

On construit les autres fonctions de lagrange par un raisonnement similaire. Cette façon de procéder est tout-à-fait générale et fonctionne pour les polynômes de degré quelconque. ◆

## C.2 Interpolation en dimension 2

Contrairement au cas unidimensionnel, nous avons ici le choix de la forme géométrique des éléments. Le plus souvent, on utilise des triangles ou des quadrilatères. Les espaces de polynômes correspondants sont légèrement différents.

### C.2.1 Interpolation sur les triangles

Voyons en premier lieu le cas des triangles. L'espace de polynômes le plus fréquemment utilisé est l'espace  $P_k$  dont on construit une base à l'aide du tableau suivant :

$$\begin{array}{l}
 P_0 : \\
 P_1 : \\
 P_2 : \\
 P_3 :
 \end{array}
 \left|
 \begin{array}{cccc}
 & & & 1 \\
 & & \xi & \eta \\
 \xi^2 & & \xi\eta & \eta^2 \\
 \xi^3 & \xi^2\eta & \xi\eta^2 & \eta^3
 \end{array}
 \right.$$

et ainsi de suite.

Pour déterminer uniquement un polynôme de degré  $k$ , on doit choisir  $N$  points d'interpolation distincts. Le choix des points d'interpolation est dicté par les propriétés de la fonction que l'on souhaite interpoler.

#### Polynômes de degré 1

Commençons par l'exemple le plus simple. Construisons les fonctions d'interpolation linéaires ( $P_1$ ) sur l'élément de référence. Du triangle précédent, la base de monômes est  $1, \xi, \eta$  de sorte que tout polynôme de  $P_1$  peut s'écrire :

$$p_1(\xi, \eta) = a_1 + a_2\xi + a_3\eta$$

La dimension de cet espace étant 3, quoi de plus naturel que de choisir les 3 sommets du triangle  $\xi_1 = (0, 0)$ ,  $\xi_2 = (1, 0)$  et  $\xi_3 = (0, 1)$  (voir la figure 6.2). On va construire, suivant la démarche utilisée en dimension 1, des fonctions d'interpolation  $L_i(\xi, \eta)$  de sorte que :

$$L_j(\xi_i) = L_j(\xi_i, \eta_i) = \delta_i^j$$

Pour la première fonction d'interpolation, le système C.2 devient :

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}
 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}
 =
 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

ce qui donne immédiatement :

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}
 =
 \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

d'où :

$$L_1(\xi_i, \eta_i) = 1 - \xi - \eta$$

de manière similaire, on trouve les fonctions  $L_2(\xi, \eta)$  et  $L_3(\xi, \eta)$ . Le tableau C.3 résume la situation. Puisqu'elles sont également utiles pour l'évaluation des systèmes élémentaires, nous avons également indiqué les dérivées partielles des polynômes de Lagrange.

#### ◆ Remarque 3.2

Polynômes de Lagrange $P_1$ (2D)			
$i$	$L_i(\xi, \eta)$	$\frac{\partial L_i}{\partial \xi}$	$\frac{\partial L_i}{\partial \eta}$
1	$1 - \xi - \eta$	-1	-1
2	$\xi$	1	0
3	$\eta$	0	1

TABLE C.3 – Polynômes de Lagrange  $P_1$  sur le triangle de référence

On pourrait également choisir comme points d'interpolation les 3 milieux de côtés. On obtient les fonctions de Lagrange dites  $P_1$  non conformes. La particularité de cette construction est que l'approximation par éléments finis qui en résulte n'est pas continue aux interfaces des éléments et n'appartient donc pas à l'espace  $H^1(\Omega)$ . ♦

### Polynômes de degré 2

Une base de l'espace  $P_2$  est constituée des polynômes  $1, \xi, \eta, \xi^2, \xi\eta$  et  $\eta^2$  ce qui constitue un espace de dimension 6 (voir la figure 6.2). Dans ce cas, on trouve le système :

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1/2 & 0 & 1/4 & 0 & 0 \\ 1 & 1/2 & 1/2 & 1/4 & 1/4 & 1/4 \\ 1 & 0 & 1/2 & 0 & 0 & 1/4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} = \mathbf{e}_i$$

En résolvant ces 6 systèmes linéaires, on trouve les fonctions du tableau C.4. Pour simplifier les expressions, on a posé :

$$\lambda = 1 - \xi - \eta$$

Polynômes de Lagrange $P_2$ (2D)			
$i$	$L_i(\xi, \eta)$	$\frac{\partial L_i}{\partial \xi}$	$\frac{\partial L_i}{\partial \eta}$
1	$-\lambda(1 - 2\lambda)$	$1 - 4\lambda$	$1 - 4\lambda$
2	$-\xi(1 - 2\xi)$	$-1 + 4\xi$	0
3	$-\eta(1 - 2\eta)$	0	$-1 + 4\eta$
4	$4\xi\lambda$	$4(\lambda - \xi)$	$-4\xi$
5	$4\xi\eta$	$4\eta$	$4\xi$
6	$4\eta\lambda$	$-4\eta$	$4(\lambda - \eta)$

TABLE C.4 – Polynômes de Lagrange  $P_2$  sur le triangle de référence

◆ **Remarque 3.3**

On peut procéder différemment et éviter complètement la résolution d'un système linéaire. L'idée est de se servir de polynômes de degré 1 et d'annuler la fonction  $L_i(\xi, \eta)$  aux endroits appropriés. Pour illustrer ce processus, regardons le 3<sup>e</sup> noeud (de coordonnée  $(0, 1)$ ) de cet élément. La fonction  $L_3(\xi, \eta)$  doit s'annuler aux autres noeuds. On peut dans un premier temps introduire le facteur  $\eta$  (qui s'annule aux noeuds 1, 2 et 4) et le facteur  $\eta - 1/2$  qui à son tour s'annule aux noeuds 5 et 6. La fonction résultante  $\eta(\eta - 1/2)$  s'annule à tous les noeuds sauf le troisième. Par contre, à ce noeud, elle vaut  $1/2$  ce qui bien sûr est différent de 1. Il suffit alors de diviser par cette valeur et on obtient :

$$L_3(\xi, \eta) = \frac{\eta(\eta - 1/2)}{1/2} = -\eta(1 - 2\eta)$$

qui est ce que l'on avait trouvé. Ce raisonnement peut s'appliquer en toutes dimensions et on y recourra fréquemment. ◆

◆ **Remarque 3.4**

On peut aussi construire des fonctions d'interpolation dites hiérarchiques. Cela consiste à utiliser des fonctions linéaires pour les 3 sommets de l'élément et des fonctions quadratiques sur les milieux de côtés. On obtient ainsi le tableau C.5.

Ceci constitue bien également un base de  $P_2$  mais ces fonctions ne vérifient pas la condition C.1. Si on pose ensuite :

$$u(\xi, \eta) = \sum_{j=1}^6 u_j L_j(\xi, \eta)$$

l'interprétation des inconnues  $u_j$  n'est plus la valeur de  $u(x)$  au noeud correspondant. En effet :

$$u(\xi_1) = u(0, 0) = u_1$$

Polynômes de Lagrange $P_2$ hiérarchique (2D)			
$i$	$L_i(\xi, \eta)$	$\frac{\partial L_i}{\partial \xi}$	$\frac{\partial L_i}{\partial \eta}$
1	$1 - \xi - \eta = \lambda$	-1	-1
2	$\xi$	+1	0
3	$\eta$	0	+1
4	$4\xi\lambda$	$4(1 - 2\xi - \eta)$	$-4\xi$
5	$4\xi\eta$	$4\eta$	$4\xi$
6	$4\eta\lambda$	$-4\eta$	$4(1 - \xi - 2\eta)$

TABLE C.5 – Polynômes de Lagrange  $P_2$  sur le triangle de référence

et c'est également le cas pour les 2 autres sommets. Par contre :

$$u(\boldsymbol{\xi}_4) = u(1/2, 0) = u_4 + \frac{(u_1 + u_2)}{2}$$

de sorte que le degré de liberté associé au quatrième noeud s'interprète par la relation :

$$u_4 = u(\boldsymbol{\xi}_4) - \frac{(u_1 + u_2)}{2} = u(\boldsymbol{\xi}_4) - \frac{(u(\boldsymbol{\xi}_1) + u(\boldsymbol{\xi}_2))}{2}$$

et l'interprétation des inconnues (éventuellement des degrés de liberté) est plus délicate. Il faudra être prudent lors de l'imposition des conditions aux limites essentielles si on utilise de telles bases.



### C.2.2 Sur les quadrilatères

Nous utiliserons l'espace  $Q_k$  sur l'élément de référence  $[-1, 1]^2$ . Il est facile de construire une base pour cet espace puisqu'il suffit de faire le produit cartésien des ensembles  $1, \xi, \xi^2, \dots, \xi^k$  et  $1, \eta, \eta^2, \dots, \eta^k$ .

#### Polynôme de degré 1 ( $Q_1$ )

Le cas le plus simple est bien sûr  $Q_1$  dont la dimension est 4. Une base possible est  $1, \xi, \eta, \xi\eta$ . Il nous faut donc 4 points d'interpolation et les 4 coins semblent un choix idéal (voir la figure 6.4). On obtient les 4 fonctions de Lagrange toujours de la même façon et nous nous limiterons à les lister.

Polynômes de Lagrange $Q_1$ (2D)			
$i$	$L_i(\xi, \eta)$	$\frac{\partial L_i}{\partial \xi}$	$\frac{\partial L_i}{\partial \eta}$
1	$\frac{1}{4}(1 + \xi)(1 + \eta)$	$\frac{1}{4}(1 + \eta)$	$\frac{1}{4}(1 + \xi)$
2	$\frac{1}{4}(1 - \xi)(1 + \eta)$	$-\frac{1}{4}(1 + \eta)$	$\frac{1}{4}(1 - \xi)$
3	$\frac{1}{4}(1 - \xi)(1 - \eta)$	$-\frac{1}{4}(1 - \eta)$	$-\frac{1}{4}(1 - \xi)$
4	$\frac{1}{4}(1 + \xi)(1 - \eta)$	$\frac{1}{4}(1 - \eta)$	$-\frac{1}{4}(1 + \xi)$

TABLE C.6 – Polynômes de Lagrange  $Q_1$  sur le carré de référence

◆ **Remarque 3.5**

Tout comme pour le cas triangulaire, on pourrait penser choisir comme points d'interpolation les 4 milieux de côtés soit les points  $(0, -1)$ ,  $(1, 0)$ ,  $(0, 1)$  et  $(-1, 0)$ . Ce qui semble naturel ne fonctionne pas toujours. En effet, en se servant encore ici de la base  $1, \xi, \eta$  et  $\xi\eta$ , la matrice du système C.2 s'écrit :

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

et est singulière. La solution n'est donc pas unique et on peut facilement le constater puisque les fonctions  $0$  et  $\xi\eta$  sont toutes deux des fonctions de  $Q_1$  qui prennent les mêmes valeurs ( $0$ ) aux milieux des côtés. On ne peut donc pas choisir ces points d'interpolation pour construire des approximations non conformes, contrairement à ce que nous avons fait sur les triangles. ◆

**Polynôme de degré 2 ( $Q_2$ )**

La dimension de l'espace  $Q_2$  est 9 et il semble naturel de choisir les points illustrés sur la figure 6.4. On construit ainsi les polynômes de Lagrange du tableau C.7.

## C.3 Interpolation en dimension 3

### C.3.1 Sur les tétraèdres

En dimension 3, on utilise les monômes suivants pour établir la base de l'approximation :

$$\begin{array}{l} P_0 : \\ P_1 : \\ P_2 : \\ P_3 : \end{array} \left| \begin{array}{l} 1 \\ \xi, \eta, \zeta, \\ \xi^2, \eta^2, \zeta^2, \xi\eta, \xi\zeta, \eta\zeta, \\ \xi^3, \eta^3, \zeta^3, \xi^2\eta, \xi^2\zeta, \eta^2\zeta, \eta^2\xi, \zeta^2\xi, \zeta^2\eta, \xi\eta\zeta, \end{array} \right.$$

De manière similaire à celle présentée dans le cas bidimensionnel (voir la figure 6.6), on montre aisément que pour les fonctions de Lagrange de degré 1, on a le tableau C.8.

Polynômes de Lagrange $Q_2$ (2D)			
$i$	$L_i(\xi, \eta)$	$\frac{\partial L_i}{\partial \xi}$	$\frac{\partial L_i}{\partial \eta}$
1	$\frac{1}{4}(1 + \xi)\xi(1 + \eta)\eta$	$\frac{1}{4}(1 + 2\xi)(1 + \eta)\eta$	$\frac{1}{4}(1 + \xi)\xi(1 + 2\eta)$
2	$-\frac{1}{4}(1 - \xi)\xi(1 + \eta)\eta$	$-\frac{1}{4}(1 - 2\xi)(1 + \eta)\eta$	$-\frac{1}{4}(1 - \xi)\xi(1 + 2\eta)$
3	$\frac{1}{4}(1 - \xi)\xi(1 - \eta)\eta$	$\frac{1}{4}(1 - 2\xi)(1 - \eta)\eta$	$\frac{1}{4}(1 - \xi)\xi(1 - 2\eta)$
4	$-\frac{1}{4}(1 + \xi)\xi(1 - \eta)\eta$	$-\frac{1}{4}(1 + 2\xi)(1 - \eta)\eta$	$-\frac{1}{4}(1 + \xi)\xi(1 - 2\eta)$
5	$\frac{1}{2}(1 - \xi^2)\eta(1 + \eta)$	$-\xi\eta(1 + \eta)$	$\frac{1}{2}(1 - \xi^2)(1 + 2\eta)$
6	$-\frac{1}{2}\xi(1 - \xi)(1 - \eta^2)$	$-\frac{1}{2}(1 - 2\xi)(1 - \eta^2)$	$\xi(1 - \xi)\eta$
7	$-\frac{1}{2}(1 - \xi^2)(1 - \eta)\eta$	$\xi\eta(1 - \eta)$	$-\frac{1}{2}(1 - \xi^2)(1 - 2\eta)$
8	$\frac{1}{2}(1 + \xi)\xi(1 - \eta^2)$	$\frac{1}{2}(1 + 2\xi)(1 - \eta^2)$	$(1 + \xi)\xi\eta$
9	$\xi(1 - \xi^2)(1 - \eta^2)$	$-2\xi(1 - \eta^2)$	$-2(1 - \xi^2)\eta$

TABLE C.7 – Polynômes de Lagrange  $Q_2$  sur le carré de référence

<b>Polynômes de Lagrange <math>P_1</math> (3D)</b>				
$i$	$L_i(\xi, \eta)$	$\frac{\partial L_i}{\partial \xi}$	$\frac{\partial L_i}{\partial \eta}$	$\frac{\partial L_i}{\partial \zeta}$
1	$1 - \xi - \eta - \zeta$	-1	-1	-1
2	$\xi$	1	0	0
3	$\eta$	0	1	0
4	$\zeta$	0	0	1

TABLE C.8 – Polynômes de Lagrange  $P_1$  sur le tétraèdre de référence

## Annexe D

# Intégration numérique

L'intégration numérique est une composante essentielle de toute méthode d'éléments finis. S'il est toujours préférable d'utiliser l'intégration exacte lorsque cela est possible, on doit toutefois recourir fréquemment à l'intégration numérique si on souhaite développer des méthodes d'éléments finis relativement générales.

Tout comme pour l'interpolation de Lagrange, nous ferons le développement des différentes formules d'intégration numérique sur les éléments de référence puisque c'est là que sont effectuées toutes les intégrales en éléments finis.

### D.1 En dimension 1

Nous choisissons l'intervalle d'intégration  $[-1, 1]$  qui est l'élément de référence pour les problèmes en dimension 1. Nous avons vu au chapitre 5 comment passer d'un élément quelconque à cet élément de référence à l'aide du changement de variable 5.6.

On cherche à évaluer une expression de la forme :

$$I = \int_{-1}^1 g(\xi) d\xi \quad (\text{D.1})$$

En éléments finis, la fonction  $g(\xi)$  fait intervenir les fonctions d'interpolation  $\hat{\psi}_i(\xi)$  et leurs dérivées ainsi que les propriétés physiques du problème. Par exemple, on doit évaluer :

$$\int_{-1}^1 p \left( \frac{(x_1^K + x_2^K) + h^K \xi}{2} \right) \hat{\psi}_j(\xi) \hat{\psi}_i(\xi) \frac{h^K}{2} d\xi$$

qui est bien de la forme D.1. Dans la plupart des programmes d'éléments finis, on utilise les quadratures de Gauss qui consistent à approcher l'intégrale D.1 par une expression de la forme :

$$\int_{-1}^1 g(\xi) d\xi \simeq \sum_{i=1}^{m_G} w_i g(\xi_i) \quad (\text{D.2})$$

qui soit la plus précise possible.

**Définition 4.1**

L'expression D.2 est appelée *quadrature de Gauss à  $m_G$  points*. Les  $\xi_i$  sont appelés *points d'intégration*, tandis que les coefficients  $w_i$  sont les *poids d'intégration*.

On choisit les points et les poids d'intégration de façon à ce que la quadrature D.2 soit exacte pour les polynômes de degré le plus élevé possible. De toute évidence, les points d'intégration  $\xi_i$  doivent tous être distincts les uns des autres et les poids d'intégration doivent être non nuls.

★ **Théorème 4.1**

La quadrature de Gauss à  $m_G$  points D.2 est exacte pour les polynômes de degré  $(2m_G - 1)$  et le terme d'erreur de cette approximation est donné par :

$$\frac{2^{2m_G+1}(m_G!)^4}{(2m_G+1)((2m_G)!)^3} g^{(2m_G)}(\mu) \quad \text{où } \mu \in [-1, 1] \quad \diamond \quad (\text{D.3})$$

★

La table D.1 résume les premières quadratures de Gauss en dimension 1 (voir Fortin, réf. [17]). La dernière colonne de cette table fournit le degré des polynômes pour lesquels la quadrature de Gauss est exacte et qui vaut  $2m_G - 1$ . C'est ce que l'on appelle le *degré de précision* de la formule de quadrature. En pratique, on choisit le nombre de points de Gauss  $m_G$  en fonction des intégrales que l'on doit évaluer.

► **Exemple 4.1**

On veut évaluer :

$$I = \int_{-1}^{+1} \xi^2 + \xi \, dt$$

Il s'agit ici d'un problème très simple mais qui illustre bien ce qui peut se produire. La fonction à intégrer est polynômiale et une quadrature de Gauss adéquate permettra d'évaluer l'intégrale exactement. Puisqu'il s'agit d'un polynôme de degré 2, la quadrature de Gauss à 2 points ( $m_G = 2$ ) suffira puisqu'elle est exacte pour des polynômes de degré jusqu'à 3 ( $2m_G - 1$ ). On a en effet :

$$\begin{aligned} I &= w_1(\xi_1^2 + \xi_1) + w_2(\xi_2^2 + \xi_2) \\ &= 1 \left( (-0,577\,350\,262\,918\,9625)^2 + (-0,577\,350\,262\,918\,9625) \right) \\ &\quad + 1 \left( (+0,577\,350\,262\,918\,9625)^2 + (+0,577\,350\,262\,918\,9625) \right) \\ &= 2/3 \end{aligned}$$

soit la valeur exacte de l'intégrale. ◀

Quadrature de Gauss (1D)			
$m_G$	Points d'intégration $\xi_i$	Poids d'intégration $w_i$	Degré de précision
1	0	2	1
2	-0,577 350 269 189 625 +0,577 350 269 189 625	1 1	3
3	-0,774 596 669 241 483 0,0 +0,774 596 669 241 483	0,555 555 555 555 556 0,888 888 888 888 889 0,555 555 555 555 556	5
4	-0,861 136 311 594 052 -0,339 981 043 584 856 +0,339 981 043 584 856 +0,861 136 311 594 052	0,347 854 845 137 454 0,652 145 154 862 545 0,652 145 154 862 545 0,347 854 845 137 454	7
5	-0,906 179 845 938 664 -0,538 469 310 105 683 0,0 +0,538 469 310 105 683 +0,906 179 845 938 664	0,236 926 885 056 189 0,478 628 670 499 365 0,568 888 889 888 889 0,478 628 670 499 365 0,236 926 885 056 189	9

TABLE D.1 – Intégration numérique sur l'intervalle de référence

## ► Exemple 4.2

On a :

$$\int_0^{\frac{\pi}{2}} \sin x dx = \frac{\pi}{2} \int_{-1}^1 \sin\left(\frac{\pi(\xi+1)}{4}\right) d\xi$$

La quadrature de Gauss à 2 points donne l'approximation :

$$\begin{aligned} \int_0^{\frac{\pi}{2}} \sin x dx &\simeq \frac{\pi}{4} \left( \sin\left(\frac{\pi(\xi_1+1)}{4}\right) + \sin\left(\frac{\pi(\xi_2+1)}{4}\right) \right) \\ &= \frac{\pi}{4} (\sin(0,331\,948\,322) + \sin(1,238\,848\,005)) \\ &= 0,998\,472\,614 \end{aligned}$$

Si on tient compte du fait que l'on a évalué la fonction  $\sin x$  en seulement deux points, ce résultat

est d'une précision remarquable. Par ailleurs, la formule à trois points donne :

$$\begin{aligned}
 \int_0^{\frac{\pi}{2}} \sin x dx &\simeq \frac{\pi}{4} \left( w_1 \sin \left( \frac{\pi(\xi_1 + 1)}{4} \right) + w_2 \sin \left( \frac{\pi(\xi_2 + 1)}{4} \right) \right. \\
 &\quad \left. + w_3 \sin \left( \frac{\pi(\xi_3 + 1)}{4} \right) \right) \\
 &= \frac{\pi}{4} \left( (0,555\ 555\ 556) \sin(0,177\ 031\ 362) \right. \\
 &\quad + (0,888\ 888\ 889) \sin(0,785\ 398\ 164) \\
 &\quad \left. + (0,555\ 555\ 556) \sin(1,774\ 596\ 669) \right) \\
 &= 1,000\ 008\ 1821
 \end{aligned}$$

◀

## D.2 En dimension 2 ou 3

### D.2.1 Sur les quadrilatères

C'est le cas le plus simple puisque l'on peut utiliser directement les techniques d'intégration numériques développées en dimension 1. Puisque dans ce cas nous avons choisi le carré  $[-1, 1]^2$  comme élément de référence, il suffit de constater que :

$$\int_{-1}^1 \left( \int_{-1}^1 g(\xi, \eta) d\xi \right) d\eta \simeq \int_{-1}^1 \left( \sum_{i=1}^{m_G} w_i g(\xi_i, \eta) \right) d\eta = \sum_{i=1}^{m_G} w_i \int_{-1}^1 g(\xi_i, \eta) d\eta \simeq \sum_{i=1}^{m_G} \sum_{j=1}^{m_G} w_i w_j g(\xi_i, \eta_j) \quad (\text{D.4})$$

En quelque sorte, l'intégration numérique en dimension 2 sur le carré  $[-1, 1]^2$  consiste donc à utiliser les formules obtenues en dimension 1 pour chacune des variables. Il en est de même en dimension 3 sur le cube de référence  $[-1, 1]^3$  et on obtient de manière similaire :

$$\int_{-1}^1 \int_{-1}^1 \int_{-1}^1 g(\xi, \eta, \zeta) d\xi d\eta d\zeta \simeq \sum_{i=1}^{m_G} \sum_{j=1}^{m_G} \sum_{k=1}^{m_G} w_i w_j w_k g(\xi_i, \eta_j, \zeta_k) \quad (\text{D.5})$$

#### ◆ Remarque 4.1

Il n'y a donc dans ce cas nul besoin de développer des méthodes d'intégration numérique particulière. Si la fonction à intégrer est un polynôme appartenant à l'espace  $Q_k$ , c'est-à-dire un polynôme de

degré  $k$  en chacune des variables d'espace, il suffit de choisir une quadrature de Gauss en dimension 1 qui soit exacte pour les polynômes de degré  $k$  et utiliser les formules [D.4](#) ou [D.5](#). ♦

### D.2.2 Sur les triangles

Les formules d'intégration numériques sont plus difficiles à obtenir sur les triangles. On ne peut plus utiliser directement les formules en dimension 1. Le tableau [D.2](#) fournit quelques unes des quadratures dites de Hammer.

On passe d'un triangle quelconque au triangle de référence à l'aide de la transformation [6.5](#). On utilise ensuite la relation :

$$\int_0^1 \int_0^{1-\xi} g(\xi, \eta) d\eta d\xi = \int_0^1 \int_0^{1-\eta} g(\xi, \eta) d\xi d\eta \simeq \sum_i w_i g(\xi_i, \eta_i)$$

où les  $w_i$ ,  $\xi_i$  et  $\eta_i$  sont donnés dans la table [D.2](#).

### D.2.3 Sur les tétraèdres

Sur les tétraèdres, il faut généralement utiliser encore plus de points d'intégration. La table [D.3](#) donne quelques unes des formules de Hammer en dimension 3.

Quadrature de Hammer (2D)				
	Coordonnées des points d'intégration		Poids d'intégration	Degré de précision
$m_G$	$\xi_i$	$\eta_i$	$w_i$	
1	+0,333 333 333 333 3333	+0,333 333 333 333 3333	+0,500 000 000 000 0000	1
3	+0,666 666 666 666 6667 +0,166 666 666 666 6667 +0,166 666 666 666 6667	+0,166 666 666 666 6667 +0,666 666 666 666 6667 +0,166 666 666 666 6667	+0,166 666 666 666 6667 +0,166 666 666 666 6667 +0,166 666 666 666 6667	2
4	+0,333 333 333 333 3333 +0,200 000 000 000 0000 +0,200 000 000 000 0000 +0,600 000 000 000 0000	+0,333 333 333 333 3333 +0,200 000 000 000 0000 +0,600 000 000 000 0000 +0,200 000 000 000 0000	-0,281 250 000 000 0000 +0,260 416 666 666 6667 +0,260 416 666 666 6667 +0,260 416 666 666 6667	3
6	+0,108 103 018 168 070 +0,445 948 490 915 965 +0,445 948 490 915 965 +0,816 847 572 980 459 +0,091 576 213 509 771 +0,091 576 213 509 771	+0,445 948 490 915 965 +0,108 103 018 168 070 +0,445 948 490 915 965 +0,091 576 213 509 771 +0,816 847 572 980 459 +0,091 576 213 509 771	+0,111 690 794 839 0055 +0,111 690 794 839 0055 +0,111 690 794 839 0055 +0,054 975 871 827 6610 +0,054 975 871 827 6610 +0,054 975 871 827 6610	4
12	+0,873 821 971 016 996 +0,063 089 014 491 502 +0,063 089 014 491 502 +0,501 426 509 658 179 +0,249 286 745 170 910 +0,249 286 745 170 910 +0,249 286 745 170 910 +0,636 502 499 121 399 +0,636 502 499 121 399 +0,310 352 451 033 785 +0,310 352 451 033 785 +0,053 145 049 844 816 +0,053 145 049 844 816	+0,063 089 014 491 502 +0,873 821 971 016 996 +0,063 089 014 491 502 +0,249 286 745 170 910 +0,501 426 509 658 179 +0,249 286 745 170 910 +0,249 286 745 170 910 +0,310 352 451 033 785 +0,053 145 049 844 816 +0,636 502 499 121 399 +0,310 352 451 033 785 +0,053 145 049 844 816 +0,310 352 451 033 785 +0,636 502 499 121 399	+0,025 422 453 185 1035 +0,025 422 453 185 1035 +0,025 422 453 185 1035 +0,058 393 137 863 1895 +0,058 393 137 863 1895 +0,058 393 137 863 1895 +0,058 393 137 863 1895 +0,041 425 537 809 1870 +0,041 425 537 809 1870	6
16	+0,333 333 333 333 333 +0,081 414 823 414 454 +0,459 292 588 292 723 +0,459 292 588 292 723 +0,658 861 384 496 480 +0,170 569 307 751 760 +0,170 569 307 751 760 +0,898 905 543 365 938 +0,050 547 228 317 031 +0,050 547 228 317 031 +0,008 394 777 409 958 +0,008 394 777 409 958 +0,263 112 829 634 638 +0,263 112 829 634 638 +0,728 492 392 955 404 +0,728 492 392 955 404	+0,333 333 333 333 333 +0,459 292 588 292 723 +0,081 414 823 414 454 +0,459 292 588 292 723 +0,170 569 307 751 760 +0,658 861 384 496 480 +0,170 569 307 751 760 +0,050 547 228 317 031 +0,898 905 543 365 938 +0,050 547 228 317 031 +0,050 547 228 317 031 +0,263 112 829 634 638 +0,728 492 392 955 404 +0,008 394 777 409 958 +0,728 492 392 955 404 +0,008 394 777 409 958 +0,263 112 829 634 638	+0,072 157 803 838 8935 +0,047 545 817 133 6425 +0,047 545 817 133 6425 +0,047 545 817 133 6425 +0,051 608 685 267 3590 +0,051 608 685 267 3590 +0,051 608 685 267 3590 +0,016 229 248 811 5990 +0,016 229 248 811 5990 +0,016 229 248 811 5990 +0,013 615 157 087 2175 +0,013 615 157 087 2175	8

TABLE D.2 – Intégration numérique sur le triangle de référence

<b>Quadrature de Hammer (3D)</b>			
Coordonnées des points d'intégration			Points d'intégration
$\xi_i$	$\eta_i$	$\zeta_i$	$w_i$
Formule à 1 point, degré de précision = 1			
+0,250 000 000 000 000	+0,250 000 000 000 000	+0,250 000 000 000 000	+0,166 666 666 666 667
Formule à 4 points, degré de précision = 2			
+0,138 196 601 125 0105	+0,138 196 601 125 0105	+0,138 196 601 125 0105	+0,041 666 666 666 667
+0,138 196 601 125 0105	+0,138 196 601 125 0105	+0,585 410 196 624 9685	+0,041 666 666 666 667
+0,138 196 601 125 0105	+0,585 410 196 624 9685	+0,138 196 601 125 0105	+0,041 666 666 666 667
+0,585 410 196 624 9685	+0,138 196 601 125 0105	+0,138 196 601 125 0105	+0,041 666 666 666 667
Formule à 16 points, degré de précision = 4			
+0,771 642 902 067 2371	+0,076 119 032 644 254 30	+0,076 119 032 644 254 30	+0,008 395 632 350 020 469
+0,076 119 032 644 254 30	+0,771 642 902 067 2371	+0,076 119 032 644 254 30	+0,008 395 632 350 020 469
+0,076 119 032 644 254 30	+0,076 119 032 644 254 30	+0,771 642 902 067 2371	+0,008 395 632 350 020 469
+0,076 119 032 644 254 30	+0,076 119 032 644 254 30	+0,076 119 032 644 254 30	+0,008 395 632 350 020 469
+0,119 700 527 797 8019	+0,071 831 645 267 669 25	+0,404 233 913 467 2644	+0,011 090 344 772 215 398
+0,119 700 527 797 8019	+0,404 233 913 467 2644	+0,071 831 645 267 669 25	+0,011 090 344 772 215 398
+0,119 700 527 797 8019	+0,404 233 913 467 2644	+0,404 233 913 467 2644	+0,011 090 344 772 215 398
+0,071 831 645 267 669 25	+0,119 700 527 797 8019	+0,404 233 913 467 2644	+0,011 090 344 772 215 398
+0,071 831 645 267 669 25	+0,404 233 913 467 2644	+0,119 700 527 797 8019	+0,011 090 344 772 215 398
+0,071 831 645 267 669 25	+0,404 233 913 467 2644	+0,404 233 913 467 2644	+0,011 090 344 772 215 398
+0,404 233 913 467 2644	+0,119 700 527 797 8019	+0,071 831 645 267 669 25	+0,011 090 344 772 215 398
+0,404 233 913 467 2644	+0,119 700 527 797 8019	+0,404 233 913 467 2644	+0,011 090 344 772 215 398
+0,404 233 913 467 2644	+0,071 831 645 267 669 25	+0,119 700 527 797 8019	+0,011 090 344 772 215 398
+0,404 233 913 467 2644	+0,071 831 645 267 669 25	+0,404 233 913 467 2644	+0,011 090 344 772 215 398
+0,404 233 913 467 2644	+0,404 233 913 467 2644	+0,119 700 527 797 8019	+0,011 090 344 772 215 398
+0,404 233 913 467 2644	+0,404 233 913 467 2644	+0,071 831 645 267 669 25	+0,011 090 344 772 215 398

TABLE D.3 – Intégration numérique sur le tétraèdre de référence



# Réponses aux exercices

## Réponses aux exercices du chapitre 2

- a)  $T'_f = \delta_0 - \delta_1$ ;  $T''_f = \delta'_0 - \delta'_1$ ;  
b)  $T'_f = T_g$  où :

$$g(x) = \begin{cases} -1 & \text{si } x < 0 \\ \cos(x) & \text{si } x > 0 \end{cases}$$

(la fonction  $f(x)$  est continue en  $x = 0$ ).

$T''_f = T'_g = T_h + 2\delta_0$  où :

$$h(x) = \begin{cases} 0 & \text{si } x < 0 \\ -\sin(x) & \text{si } x > 0 \end{cases}$$

(la fonction  $g(x)$  a un saut de hauteur 2 en  $x = 0$ ).

c)  $T'_f = T_g + \delta_0$  où :

$$g(x) = \begin{cases} -1 & \text{si } x < 0 \\ -\sin(x) & \text{si } x > 0 \end{cases}$$

(la fonction  $f(x)$  a un saut de hauteur 1 en  $x = 0$ ).

$T''_f = T'_g + \delta'_0 = T_h + \delta_0 + \delta'_0$  où :

$$h(x) = \begin{cases} 0 & \text{si } x < 0 \\ -\cos(x) & \text{si } x > 0 \end{cases}$$

(la fonction  $g(x)$  possède un saut de hauteur 1 en  $x = 0$ ).

- Dans  $L^2(] - 1, 1[)$  : a, b et c ; Dans  $H^1(] - 1, 1[)$  : b seulement ; Dans  $H^2(] - 1, 1[)$  : aucune.
- La fonction  $u$  est continue sur l'axe  $y = x$ . De plus,  $\nabla u_1 = (-2, 2)$ ,  $\mathbf{n}_1 = (\sqrt{2}/2, -\sqrt{2}/2)$ ,  $\nabla u_2 = (2x, -2y)$ , et  $\mathbf{n}_2 = (-\sqrt{2}/2, \sqrt{2}/2)$ . Le saut de la dérivée normale est donc :  $k_2 \nabla u_2 \cdot \mathbf{n}_1 - k_1 \nabla u_1 \cdot \mathbf{n}_1 = x + y + 4$ . La divergence est donc :

$$\begin{cases} 0 & \text{si } x < y \\ 0 & \text{si } x > y \end{cases}$$

plus une simple couche le long de l'axe  $y = x$  d'intensité  $\mathbf{s} \cdot \mathbf{n} = x + y + 4$ .

4. La fonction  $f(x) = 1$  est dans  $L^1_{loc}(\Omega)$ . Il suffit de remarquer que l'on intègre sur un compact (fermé et borné); Pour la fonction  $f(x) = \ln x$ , la difficulté vient du point  $x = 0$ . Or, un compact de  $]0, 1[$  ne peut pas contenir ce point et la fonction sera intégrable.
5.  $|x|$  est dans  $H^1(]-1, 1[)$  car elle est de carré sommable et sa dérivée au sens des distributions est  $T_g$  où :

$$g(x) = \begin{cases} -1 & \text{si } x < 0 \\ +1 & \text{si } x > 0 \end{cases}$$

qui est aussi de carré sommable). Par contre, elle n'est pas dans  $H^2(]-1, 1[)$  puisque la dérivée seconde fait apparaître une distribution de Dirac en  $x = 0$ .

6. Pour  $f(x) = x(1-x)$ ,  $f(0) = f(1) = 0$  mais  $f'(0) = 1 \neq 0$  et cette fonction n'est pas dans  $H^2_0(]0, 1[)$ . Par contre, on vérifie facilement que la fonction  $f(x) = x^2(1-x)^2$  est bien dans  $H^2_0(]0, 1[)$ .
7.  $x^r$  est dans  $L^2(]0, 1[)$  pour  $r > -1/2$ ; dans  $H^1(]0, 1[)$  pour  $r > 1/2$  et  $r = 0$ ; dans  $H^2(]0, 1[)$  pour  $r > 3/2$ ,  $r = 0$  et  $r = 1$ .

### Réponses aux exercices du chapitre 3

1. La symétrie des 4 formes bilinéaires découle de la commutativité du produit de deux fonctions et du produit scalaire de deux vecteurs.

a)

$$|a(u, w)| \leq p_2 \int_{\Omega} u w \, dv \leq p_2 \|u\|_{0,\Omega} \|w\|_{0,\Omega}$$

$$|a(w, w)| \geq p_1 \int_{\Omega} w w \, dv = p_1 \|w\|_{0,\Omega}^2$$

b)

$$|a(u, w)| \leq q_2 \int_{\Omega} \nabla u \cdot \nabla w \, dv \leq q_2 \sum_{i=1}^n \int_{\Omega} \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_i} \, dv \leq nq_2 |u|_{1,\Omega} |w|_{1,\Omega}$$

$$|a(w, w)| \geq q_1 \int_{\Omega} \nabla w \cdot \nabla w \, dv = q_1 |w|_{1,\Omega}^2$$

en remarquant que  $|w|_{1,\Omega}$  est bien une norme sur  $H^1_0(\Omega)$ .

c)

$$|a(u, w)| \leq q_2 \int_{\Omega} \nabla u \cdot \nabla w \, dv \leq q_2 \sum_{i=1}^n \int_{\Omega} \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_i} \, dv \leq nq_2 |u|_{1,\Omega} |w|_{1,\Omega} \leq nq_2 \|u\|_{1,\Omega} \|w\|_{1,\Omega}$$

$$|a(w, w)| \geq q_1 \int_{\Omega} \nabla w \cdot \nabla w \, dv = q_1 |w|_{1,\Omega}^2$$

et on ne peut pas conclure car  $\|w\|_{1,\Omega}$  n'est pas une norme sur  $H^1(\Omega)$ . La forme bilinéaire n'est pas coercive sur  $H^1(\Omega)$ .

d)

$$\begin{aligned} |a(u, w)| &\leq p_2 \int_{\Omega} u w \, dv + q_2 \int_{\Omega} \nabla u \cdot \nabla w \, dv \leq p_2 \|u\|_{0,\Omega} \|w\|_{0,\Omega} + q_2 \sum_{i=1}^n \int_{\Omega} \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_i} \, dv \\ &\leq p_2 \|u\|_{0,\Omega} \|w\|_{0,\Omega} + nq_2 \|u\|_{1,\Omega} \|w\|_{1,\Omega} \leq (p_2 + nq_2) \|u\|_{1,\Omega} \|w\|_{1,\Omega} \\ |a(w, w)| &\geq p_1 \int_{\Omega} w w \, dv + q_1 \int_{\Omega} \nabla w \cdot \nabla w \, dv \geq \min(p_1, q_1) \|w\|_{1,\Omega}^2 \end{aligned}$$

2. a) On pose  $V = H_0^1(]0, 1[)$  et la formulation variationnelle est :

$$\int_0^1 \nu_2 u'(x) w'(x) \, dx = \int_0^1 f(x) w(x) \, dx \quad \forall w \in V$$

Le relèvement des conditions essentielles est nul.

b) On pose  $V = H^1(]0, 1[)$  et la formulation variationnelle est :

$$\int_0^1 \nu_1 u w + \nu_2 u'(x) w'(x) \, dx = \int_0^1 f(x) w(x) \, dx \quad \forall w \in V$$

Il n'y a pas de relèvement à faire.

c) On pose  $V = H_0^1(]0, 1[)$  et par exemple  $u_g = a + (b-a)x$  pour le relèvement. La formulation variationnelle est :

$$\int_0^1 \nu_2 \delta'_u(x) w'(x) \, dx = \int_0^1 (f(x) w(x) - \nu_2 (b-a) w'(x)) \, dx \quad \forall w \in V$$

Notons que le dernier terme de droite, qui tient compte du relèvement de la condition essentielle, s'annule pour cet exemple, mais que ce n'est pas le cas en général.

d) On pose  $V = \{w \in H^1(]0, 1[) \mid w(0) = 0\}$  et par exemple  $u_g = a$  pour le relèvement. La formulation variationnelle est :

$$\int_0^1 q(x) \delta'_u(x) w'(x) \, dx = \int_0^1 f(x) w(x) \, dx + b w(1) \quad \forall w \in V$$

Pour démontrer la continuité de  $l(w)$ , il faut utiliser la continuité de la trace au bord.

e) On pose  $V = \{w \in H^2(]0, L[) \mid w(0) = w(L) = 0\}$  et par exemple  $u_g = a + (b-a)x/L$  pour le relèvement. La formulation variationnelle est :

$$\int_0^L q(x) \delta''_u(x) w''(x) \, dx = \int_0^L f(x) w(x) \, dx + d w'(L) - c w'(0) \quad \forall w \in V$$

Pour démontrer la continuité de  $l(w)$ , il faut utiliser la continuité de la trace au bord.

f) On pose  $V = \{w \in H^1(\Omega) | w = 0 \text{ sur } \Gamma_0\}$  et on prendra un relèvement  $u_g$  tel que  $u_g = g$  sur  $\Gamma_0$ . La formulation variationnelle est :

$$\int_{\Omega} q(x) \nabla \delta_u \cdot \nabla w \, dv = \int_{\Omega} f(x) w(x) \, dx - \int_{\Omega} q(x) \nabla u_g \cdot \nabla w \, dv + \int_{\Gamma_1} h w \, ds \quad \forall w \in V$$

g) On pose  $V = H^1(\Omega)$  et il n'y a pas de relèvement à faire. La formulation variationnelle est :

$$\int_{\Omega} \nabla \delta_u \cdot \nabla w \, dv = \int_{\Omega} f(x) w(x) \, dx + \int_{\Gamma} h w \, ds \quad \forall w \in V$$

La forme bilinéaire n'est pas coercive sur  $V$  (voir exercice 1-c). On remarque également que si on a une solution  $u$ , alors  $u + c$  est aussi une solution. Il n'y a pas unicité.

h) On pose  $V = H^1(\Omega)$  et il n'y a pas de relèvement à faire. La formulation variationnelle est :

$$\int_{\Omega} \nabla u \cdot \nabla w \, dv + \int_{\Gamma} u w \, ds = \int_{\Omega} f(x) w(x) \, dx + \int_{\Gamma} h w \, ds \quad \forall w \in V$$

Pour démontrer la coercivité de la forme bilinéaire, il faudrait démontrer que  $(|w|_{1,\Omega}^2 + |w|_{0,\Gamma}^2)^{1/2}$  est une norme sur  $H^1(\Omega)$  équivalente à la norme habituelle, ce qui est le cas.

3. Les deux dernières conditions aux limites ne peuvent être imposées au même endroit. Le problème est mal posé.
4. On a  $a(u, w) = l(w)$  pour tout  $w$  dans  $V$  et donc  $a(u, u) = l(u)$ . Le résultat suit en vertu de la coercivité et de la définition de  $\|l\|$ .
5. On pose  $V = H_0^1(]0, 1[)$  et la formulation variationnelle est :

$$\int_0^1 c_1 u'(x) w'(x) + c_2 u'(x) w(x) \, dx = \int_0^1 f(x) w(x) \, dx \quad \forall w \in V$$

La continuité des formes linéaire et bilinéaire est évidente. Pour la coercivité, on a :

$$|a(w, w)| \geq c_1 \int_{\Omega} (w'(x))^2 + c_2 w'(x) w(x) \, dv = c_1 |w|_{1,\Omega}^2 + \frac{c_2}{2} w(x) \Big|_0^1 = c_1 |w|_{1,\Omega}^2$$

car le deuxième terme s'annule.

## Réponses aux exercices du chapitre 4

1.  $\phi_i(x) = (x(1-x))^i$ , pour  $i \geq 1$ .
2.  $\phi_i(x) = x^i$ , pour  $i \geq 0$  (il est important d'inclure la fonction  $\phi_0(x) = 1$  car autrement la solution serait forcément nulle en  $x = 0$ ).
3.  $u_g(x) = a + (b-a)x$  et  $\phi_i(x) = (x(1-x))^i$ , pour  $i \geq 1$ .

4.  $u_g(x) = a$  et  $\phi_i(x) = x^i$ , pour  $i \geq 1$ .
5.  $u_g(x) = a + (b-a)x/L$  et  $\phi_i(x) = (x(L-x))^i$ , pour  $i \geq 1$ .
6. Le relèvement doit vérifier les 3 premières conditions aux limites et on prendra par exemple  $u_g(x) = a + \frac{b-a}{L^2}x^2$ . Les fonctions  $\phi_i(x)$  doivent donc s'annuler en 0 et  $L$  et leur dérivée doit s'annuler en 0. On prendra par exemple  $\phi_i(x) = (x^2(L-x))^i$  pour  $i \geq 1$ .
7. Le relèvement est  $u_g(x, y) = 1$  et  $\phi_i(x, y) = (xy)^i$ .
8. Le relèvement est  $u_g(x) = (x+1)^2$  et  $\phi_i(x) = (x^2(1-x))^i$  pour  $i \geq 1$ .
9. On pose  $V = \{w \in H^1(\Omega) \mid w = 0 \text{ sur les axes } x = 1 \text{ et } y = 1\}$  et il n'y a pas de relèvement à faire. La formulation variationnelle est :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv = \int_{\Omega} q w(x) \, dx \quad \forall w \in V$$

On a donc :

$$A_{11} = \int_0^1 \int_0^1 k (-2x(1-y^2), -2y(1-x^2)) \cdot (-2x(1-y^2), -2y(1-x^2)) \, dx dy = \frac{64k}{15}$$

et :

$$F_1 = \int_0^1 \int_0^1 q(1-x^2)(1-y^2) \, dx dy = \frac{4q}{9}$$

Le coefficient de la seule fonction de Ritz sera donc  $5q/(48k)$ .

10. On pose  $V = \{w \in H^1(]0, 1[) \mid w(0) = 0\}$  et par exemple  $u_g = 1$  pour le relèvement. La formulation variationnelle est :

$$\int_0^1 q(x) \delta'_u(x) w'(x) \, dx = \int_0^1 (2 - 2x + 6x^2) w(x) \, dx - 2w(1) \quad \forall w \in V$$

On prend ensuite  $\phi_i(x) = x^i$  pour  $i \geq 1$ . On a alors :

$$A_{ij} = ij \left( \frac{1}{i+j+1} + \frac{1}{i+j-1} \right) \text{ et } F_i = \frac{2}{i+1} - \frac{2}{i+2} + \frac{6}{i+3} - 2$$

Si on prend 2 fonction de Ritz, on trouve  $c_1 = 1$  et  $c_2 = -1$  de sorte que la solution approchée est  $u(x) = 1 + x - x^2$ .

11. La matrice de la méthode de Ritz a pour terme général  $A_{ij} = a(\phi_j, \phi_i)$ . On a alors :

$$\mathbf{Ax} = \sum_{j=1}^n A_{ij} x_j = \sum_{j=1}^n a(\phi_j, \phi_i) x_j = a \left( \sum_{j=1}^n x_j \phi_j, \phi_i \right) = a(w, \phi_i) \text{ où } w(x) = \sum_{j=1}^n x_j \phi_j(x)$$

on a donc :

$$\mathbf{x}^T \mathbf{Ax} = \sum_{i=1}^n a(w, \phi_i) x_i = a(w, w) \geq \alpha \|w\|_V > 0$$

## Réponses aux exercices du chapitre 5

1. On a la situation suivante :

$$u^K(x_i^K) = \sum_{j=1}^3 u_j^K \psi_j^K(x_i^K) = \sum_{j=1}^3 u_j^K \hat{\psi}_j(\xi_i)$$

Si on se place aux noeuds géométriques, on a  $u^K(x_i^K) = u_i^K$  mais au point milieu, on a ;

$$u^K(x_3^K) = \sum_{j=1}^3 u_j^K \hat{\psi}_j(0) = \frac{u_1^K + u_2^K}{2} + u_3^K$$

et par conséquent :

$$u_3^K = u^K(x_3^K) - \frac{u_1^K + u_2^K}{2} = u^K(x_3^K) - \left( \frac{u(x_1^K) + u(x_2^K)}{2} \right)$$

2.

3. Pour la première fonction, on cherche une expression de la forme  $\tilde{\psi}_1(\xi) = a_0 + a_1\xi + a_2\xi^2 + a_3\xi^3$ . Les conditions à vérifier sont  $\tilde{\psi}_1(-1) = 1$ ,  $\tilde{\psi}_1'(-1) = 0$ ,  $\tilde{\psi}_1(+1) = 0$  et  $\tilde{\psi}_1'(+1) = 0$ . On obtient le système :

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

On trouve  $a_0 = 1/2$ ,  $a_1 = -3/4$ ,  $a_2 = 0$  et  $a_3 = 1/4$ . On a alors :

$$\tilde{\psi}_1(x) = \frac{1}{2} - \frac{3}{4}\xi + \frac{1}{4}\xi^3 = \frac{1}{4}(1 - \xi)^2(2 + \xi)$$

Les autres fonctions de base s'obtiennent de manière similaire.

4. a) Sur le premier élément :

$$u^{K_1}(x) = \sum_{j=1}^3 u_j^{K_1} \phi_j^{K_1} = \sum_{j=1}^3 u_j^{K_1} \hat{\psi}_j(\xi)$$

On a ainsi :

$$\begin{aligned} \frac{du^{K_1}}{dx}(0) &= \sum_{j=1}^3 u_j^{K_1} \frac{2}{h^{K_1}} \frac{d\hat{\psi}_j}{d\xi}(-1) \\ &= \frac{2}{h^{K_1}} \left( -0,355 \frac{d\hat{\psi}_2}{d\xi}(-1) - 0,235 \frac{d\hat{\psi}_3}{d\xi}(-1) \right) \\ &= -0,2925 \end{aligned}$$

puisque  $h^{K_1} = 2$ . On a donc  $-400 \frac{du^{K_1}}{dx}(0) = 117$ . Le calcul de la même quantité en utilisant cette fois la variable secondaire  $s_{11}^{K_1}$  donne 103,67 qui est beaucoup plus précis.

b) le relèvement serait  $2\phi_8 - 3\phi_9$  ou encore  $2\psi_1^{K_1}(x) - 3\psi_2^{K_4}(x)$ .

5. On doit évaluer la matrice élémentaire de terme général :

$$a_{ij}^K = \int_{x_1^K}^{x_2^K} c_2(\psi_j^K)'(x)\psi_i^K(x) dx = \int_{-1}^1 c_2\hat{\psi}_j'(\xi)\hat{\psi}_i(\xi) d\xi$$

et on a :

$$A^K = \frac{c_2}{2} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$$

6. La formulation variationnelle (sans faire le relèvement) est :

$$\int_0^1 (x^3 u'(x)w'(x) + x^2 u(x)w(x)) dx = \int_0^1 \frac{w(x)}{1+x} dx$$

a)  $\psi_3(\xi) = \frac{27}{16}(\xi^2 - 1)(\xi - 1/3)$ .

b) Le relèvement implicitement construit sera  $-\phi_9 + 3\phi_{10}$  (les  $\phi_i$  sont les fonctions de Ritz cubiques par élément).

c) Le degré maximal à intégrer est 8. Il faut donc 5 points de Gauss.

d)

Numéros des ddl				
Table <i>adres</i>				
Élément	Ddl #1	Ddl #2	Ddl #3	Ddl #4
1	9	3	1	2
2	3	6	4	5
3	6	10	7	8

e)  $a_{33} = a_{22}^{K_1} + a_{11}^{K_2}$  et  $a_{14} = 0$  (les degrés de liberté 1 et 4 n'étant pas connectés).

## Réponses aux exercices du chapitre 6

1.

2. Il suffit de remarquer que :

$$\int_K 1 dv = \int_{\hat{K}} J^K d\hat{v} \text{ ou encore } mes(K) = J^K mes(\hat{K})$$

si le jacobien  $J^K$  est constant et où  $mes$  désigne l'aire en dimension 2 et le volume en dimension 3.

- 3.
- 4.
- 5.
- 6.
- 7.
8. Puisque les côtés sont droits, la transformation géométrique est linéaire pour les triangles et bilinéaires pour les quadrangles. Sur un triangle, après passage à l'élément de référence, l'intégrand est de degré 3 puisque le jacobien est constant de même que les coefficients de la matrice  $B^K$  (voir la relation 6.10). Une formule de Hammer à 4 points intégrera exactement. Par contre, sur un quadrangle, le jacobien ne sera pas constant et les coefficients de  $B^K$  sont des quotients de polynômes (voir la relation 6.12). Aucune formule de quadrature de Gauss ne pourra intégrer exactement.
9.  $\hat{\psi}_4(\xi, \eta) = \frac{27}{2}\xi(1 - \xi - \eta)(\frac{2}{3} - \xi - \eta)$
- 10.
11. En inversant la transformation linéaire, on montre que :

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \frac{1}{18} \begin{bmatrix} 4 & -2 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix}$$

Le point  $(4, 3)$  correspond donc à  $(\frac{2}{9}, \frac{4}{9})$  dans l'élément de référence et :

$$u(4, 3) = \sum_{j=1}^3 u_j^K \hat{\psi}_j(\frac{2}{9}, \frac{4}{9}) = 31,66$$

de même que :

$$\begin{aligned} \frac{\partial u}{\partial x}(4, 3) &= \sum_{j=1}^3 u_j^K \left( \frac{\partial \hat{\psi}_j}{\partial \xi}(\frac{2}{9}, \frac{4}{9}) \frac{\partial \xi}{\partial x} + \frac{\partial \hat{\psi}_j}{\partial \eta}(\frac{2}{9}, \frac{4}{9}) \frac{\partial \eta}{\partial x} \right) \\ &= \sum_{j=1}^3 u_j^K \left( \frac{\partial \hat{\psi}_j}{\partial \xi}(\frac{2}{9}, \frac{4}{9}) \frac{4}{18} + \frac{\partial \hat{\psi}_j}{\partial \eta}(\frac{2}{9}, \frac{4}{9}) \frac{-1}{18} \right) = 8,66 \end{aligned}$$

## Réponses aux exercices du chapitre 8

1. La formulation variationnelle (non linéarisée) est :

$$\int_{\Omega} [k(1 + \lambda u^2)^m] \nabla u \cdot \nabla w \, dv = \int_{\Omega} f(x_1, x_2) w(x_1, x_2) \, dv + \int_{\Gamma_1} h w \, ds$$

À la première itération, on peut prendre  $u_0$  solution du problème linéaire suivant :

$$\int_{\Omega} k \nabla u_0 \cdot \nabla w \, dv = \int_{\Omega} f(x_1, x_2) w(x_1, x_2) \, dv + \int_{\Gamma_1} h w \, ds$$

obtenu en prenant  $m = 0$  ou  $\lambda = 0$  pour rendre le problème linéaire. On devra ainsi introduire un relèvement des conditions aux limites essentielles et résoudre :

$$\int_{\Omega} k \nabla \delta_u \cdot \nabla w \, dv = - \int_{\Omega} k \nabla u_g \cdot \nabla w \, dv + \int_{\Omega} f(x_1, x_2) w(x_1, x_2) \, dv + \int_{\Gamma_1} h w \, ds$$

et poser  $u_0 = u_g + \delta_u$ . Pour les itérations suivantes, on remet  $m$  ou  $\lambda$  à sa vraie valeur et on pose  $u = u_k + \delta_u$  ( $u_k$  vérifiant les conditions essentielles) et :

$$R(u, w) = \int_{\Omega} [k(1 + \lambda u^2)^m] \nabla u \cdot \nabla w \, dv - \int_{\Omega} f(x_1, x_2) w(x_1, x_2) \, dv - \int_{\Gamma_1} h w \, ds$$

On linéarise ensuite en dérivant :

$$\begin{aligned} D_u R(u_k, w)[\delta_u] &= \frac{d}{d\epsilon} [R(u_k + \epsilon \delta_u, w)] |_{\epsilon=0} \\ &= \int_{\Omega} [k(1 + \lambda u_k^2)^m] \nabla \delta_u \cdot \nabla w \, dv \\ &\quad + \int_{\Omega} [2\lambda m k u_k (1 + \lambda u_k^2)^{m-1}] \delta_u \nabla u_k \cdot \nabla w \, dv \end{aligned}$$

On résout alors à chaque itération :

$$\int_{\Omega} [k(1 + \lambda u_k^2)^m \nabla \delta_u] \cdot \nabla w \, dv + \int_{\Omega} [2\lambda m k u_k (1 + \lambda u_k^2)^{m-1}] \delta_u \nabla u_k \cdot \nabla w \, dv = -R(u_k, w)$$

et on met à jour en posant :  $u_{k+1} = u_k + \delta_u$ . On itérera ainsi jusqu'à ce que  $\|\delta_u\|_{\infty} < \epsilon$  où  $\epsilon$  est la tolérance désirée.

2. La non-linéarité provient cette fois de la condition aux limites. La formulation variationnelle non linéarisée est :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv + \int_{\Gamma_1} h(u^4 - u_{\infty}^4) w \, ds = \int_{\Omega} f(x_1, x_2) w(x_1, x_2) \, dv$$

On peut démarrer les itérations pour résoudre ce problème en résolvant le problème linéaire :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv = \int_{\Omega} f(x_1, x_2) w(x_1, x_2) \, dv$$

obtenu en posant  $h = 0$  ou mieux encore, en résolvant :

$$\int_{\Omega} k \nabla u \cdot \nabla w \, dv + \int_{\Gamma_1} h(u - u_{\infty}^4) w \, ds = \int_{\Omega} f(x_1, x_2) w(x_1, x_2) \, dv$$

qui est aussi linéaire mais plus près du véritable problème. Les itérations de la méthode de Newton ont la forme :

$$\int_{\Omega} k \nabla \delta_u \cdot \nabla w \, dv + \int_{\Gamma_1} 4h(u_k)^3 \delta_u w \, ds = -R(u_k, w)$$

où :

$$R(u_k, w) = \int_{\Omega} k \nabla u \cdot \nabla w \, dv + \int_{\Gamma_1} h(u^4 - u_{\infty}^4) w \, ds - \int_{\Omega} f(x_1, x_2) w(x_1, x_2) \, dv$$

## Réponses aux exercices du chapitre 14

- 1.
- 2.
- 3.
- 4.
- 5.
6. On a que :

$$\frac{\partial J_1}{\partial \mathbf{C}} \cdot \mathbf{C} = I_3^{-1/3} \mathbf{C} - \frac{1}{3} I_1 I_3^{-1/3} \mathbf{I}$$

de sorte que la trace est nulle. De plus :

$$\frac{\partial J_2}{\partial \mathbf{C}} \cdot \mathbf{C} = (I_1 \mathbf{C} - \mathbf{C} \cdot \mathbf{C}) I_3^{-2/3} - \frac{2}{3} I_2 I_3^{-2/3} \mathbf{I}$$

et on rappelle que  $\text{tr}(\mathbf{C} \cdot \mathbf{C}) = I_1^2 - 2I_2$ . Le résultat suit immédiatement.

# Bibliographie

- [1] Agassant, J. F., P. Avenas, J. Ph. Sergent et P. Carreau: *Polymer Processing, Principles and Modeling*. Hanser Publishers, Munich, 1991.
- [2] Arnold, D., F. Brezzi et M. Fortin: *A Stable Finite Element for the Stokes Equations*. *Calcolo*, 21 :337–344, 1984.
- [3] Bartle, R.G.: *The elements of integration*. Wiley, New York, 1966.
- [4] Bathe, K. J.: *Finite Element Procedures*. Prentice Hall, 1996.
- [5] Bathe, K. J. et T. Sussman: *A Finite Element Formulation for Nonlinear Incompressible Elastic and inelastic Analysis*. *Computers and Structures*, 26(1/2) :357–409, 1987.
- [6] Bonet, J. et D. R. Wood: *Nonlinear continuum mechanics for finite element analysis*. Cambridge, 1997.
- [7] Brezzi, F. et M. Fortin: *Mixed and Hybrid Finite Element Methods*. Springer-Verlag, 1991.
- [8] Burden, R.L. et J.D. Faires: *Numerical Analysis*. Brooks/Cole, Pacific Grove, septième édition, 2001, ISBN 0-534-38216-9.
- [9] Ciarlet, P.G.: *The Finite Element Method for Elliptic problems*. North-Holland, Amsterdam, 1986.
- [10] Courant, R. L.: *Variational Methods for the Solution of Problems of Equilibrium and Vibration*. *Bulletin of the American Mathematical Society*, 49 :1–23, 1943.
- [11] C ea, J.: *Approximation variationnelle des probl emes aux limites*. *Ann. Inst. Fourier (Grenoble)*, 14 :345–444, 1964.
- [12] Delfour, M.C. et J. P. Zol ezio: *Shapes and Geometries : Analysis, Differential Calculus, and Optimization*. *Advances in Design and Control*. SIAM, Philadelphia, 2001.
- [13] Dhatt, G. et G. Touzot: *Une pr esentation de la m ethode des  el ements finis*. Les presses de l’Universit e Laval, Qu ebec, 1981.
- [14] Duvaut, G.: *M ecanique des milieux continus*. *Collection math ematiques appliqu ees pour la ma ıtrise*. Masson, 1990.
- [15] Duvaut, G. et J.L. Lions: *Les in equations en m ecanique et en physique*. Dunod, Paris, 1972.
- [16] Engelman, M. et M. Jamnia: *Transient Flow Past a Circular Cylinder : a Benchmark Solution*. *Int. J. Numer. Meth. Fluids*, 11 :985–1000, 1990.

- [17] Fortin, A.: *Analyse numérique pour ingénieurs*. Presses Internationales Polytechnique, troisième édition, 2008.
- [18] Fortin, A., M. Jardak, R. Pierre et J. J. Gervais: *Localization of Hopf Bifurcations in Fluid Flow Problems*. Int. J. Numer. Meth. Fluids, 24 :1185–1210, 1997.
- [19] Fortin, M. et A. Fortin: *Experiments with Several Elements for Viscous Incompressible Flows*. Int. J. Num. Meth. Fluids, 5 :911–928, 1985.
- [20] Fortin, M. et R. Glowinski: *Résolution Numérique de Problèmes aux Limites par des Méthodes de Lagrangien Augmenté*. Dunod, 1983.
- [21] Gasquet, C. et P. Witomski: *Analyse de Fourier et applications*. Masson, Paris, 1990.
- [22] George, A. et J. Liu: *Computer Solution of Large Sparse Positive Definite Systems*. Prentice-Hall, 1981.
- [23] Jean, R.: *Mesure et intégration*. Les presses de l’université du Québec, Montréal, 1975.
- [24] Lucquin, B.: *Équations aux dérivées partielles et leurs approximations*. Mathématiques à l’université. Ellipses, Paris, 2004.
- [25] The MathWorks Inc.: *Matlab User’s guide*, 1996.
- [26] Medvinsky, A. B., S. V. Petrovskii, I. A. Tikhonova, H. Malchow et B. L. Li: *Spatiotemporal Complexity of Plankton and Fish Dynamics*. SIAM review, 44(3), 2002.
- [27] Ozell, B. et C. Pic: *Manuel d’utilisation du logiciel configurable de visualisation scientifique VU*. CERCA, Montréal, 1998.
- [28] Philippin, G.: *Cours d’analyse à l’usage des ingénieurs*, tome 2. Presses de l’Université de Montréal, Montréal, 1994.
- [29] Raviart, P.A. et J.M. Thomas: *Introduction à l’analyse numérique des équations aux dérivées partielles*. Masson, Paris, 1983.
- [30] Reddy, J.N.: *Applied Functional Analysis and Variational Methods in Engineering*. McGraw Hill, New York, 1986.
- [31] Reddy, J.N.: *An Introduction to the Finite Element Method*. McGraw-Hill, New York, deuxième édition, 1993.
- [32] Ritz, W: *Ueber eine neue Methode zur Losung gewisser Variations probleme der mathematischen Physik*. J. Reine Angew. Math., 135 :1–61, 1908.
- [33] Roache, P.J.: *Code verification by the method of manufactured solutions*. ASME, 124 :4–10, mars 2002.
- [34] Schwartz, L.: *Théorie des distributions*. Dunod, Paris, 1965.
- [35] Strang, G. et G. Fix: *An Analysis of the Finite Element Method*. Prentice-Hall, Englewood Cliffs, 1973.
- [36] Swokowski: *Analyse*. De Boeck Université, Bruxelles, cinquième édition, 1993.

- [37] Turner, M.J., R.W. Clough, H.C. Martin et L.C. Topp: *Stiffness and Deflection Analysis of Complex Structures*. Journal of the Aeronautical Sciences, 23 :805–882, 1956.
- [38] Youbissi, Fabien M.: *Résolution par éléments finis du problème de contact unilatéral par des méthodes d'optimisation convexe*. Thèse de doctorat, Université Laval, Québec, 2006. <http://www.theses.ulaval.ca/2006/23948/23948.pdf>.
- [39] Zienkiewicz, O. C. et R. L. Taylor: *The finite Element Method : basic formulation and linear problems*, tome 1. McGraw-Hill, 1989.

# Index

- $\mathcal{D}'(\Omega)$ , 13
- écoulement de Poiseuille, 258, 263
- élément Mini, 253
- élément de Crouzeix et Raviart, 256
- élément de Taylor-Hood, 254
- élément de référence, 87
- élément réel, 87
- éléments, 77, 79
- éléments compatibles, 253
- équation de réaction-diffusion, 208
- évolution de la masse volumique, 294
  
- algorithme de Newton, 184, 189
- allée de von Karman, 269
- assemblage, 104
  
- base hiérarchique, 93, 135, 371
- bifurcation de Hopf, 269
  
- coefficient de Poisson, 223
- coefficients de Lamé, 223, 276
- compact, 8
- complétude, 29
- condition aux limites naturelle, 50
- condition de Kuhn-Tucker, 333
- condition naturelle, 56, 141
- conditions aux limites essentielles, 51
- contact unilatéral, 333
- continuité d'une fonctionnelle, 13
- convergence, 168
  
- déviateur, 276
- degré de liberté, 80, 81
- degré de précision, 95, 151, 378
  
- degrés de liberté, 139
- deuxième tenseur de Piola-Kirchoff, 298
- différentiation amont, 212
- discontinuités de première espèce, 16
- distribution, 13
- distribution de Dirac, 13
- distribution régulière, 15
- divergence d'une distribution, 21
- divergence tangentielle, 318
- domaine de définition d'une fonctionnelle, 9
- déformation isochorique, 302
- dérivée d'une distribution, 16
- dérivée de Gâteaux, 202
  
- espace  $P_k$ , 365
- espace complet, 29
- espace de Hilbert, 29
- espace dual, 13, 44
- espace fonctionnel linéaire, 24
- espaces de Sobolev, 29
  
- flux, 344
- fonction continue par morceaux, 16
- fonction de Heaviside, 17
- fonction distance, 329
- fonction distance signée, 329
- fonction localement intégrable, 14
- fonctionnelle, 9
- fonctionnelle d'énergie, 47
- fonctionnelle linéaire, 12
- fonctionnelle énergie, 12
- fonctions d'interpolation, 86
- fonctions de Ritz, 63
- forme bilinéaire, 44

- forme bilinéaire coercive, 46
- forme bilinéaire elliptique, 46
- forme bilinéaire symétrique, 45
- forme linéaire, 43
- forme linéaire continue, 43
- formulation faible, 4
- formulation forte, 4, 5
- formulation quasi-variationnelle, 197
- formulation variationnelle, 4
- formule de Nanson, 292
- formule de Piola, 293
  
- gradient, 341
- gradient de déformation, 290
- génération de maillage, 138
  
- inéquation variationnelle, 334
- intégrale de Lebesgue, 11
- inégalité de Korn, 224
- inégalité triangulaire, 28
  
- Jacobien, 144
- Jacobien surfacique, 293
  
- largeur de bande , 139
- largeur de bande maximale, 140
- loi de Carreau-Yasuda, 262
- loi de comportement, 221
- loi de frottement de Coulomb, 337
- loi de puissance, 263
- loi puissance, 262
  
- maître-esclave, 329
- maillage, 79
- matrice bande, 110, 140
- matrice creuse, 102, 139
- matrice de rigidité, 86
- matrice définie positive, 46
- matrice en ligne de ciel, 110, 140
- matrice jacobienne, 144, 183
- matrice masse, 199
- matrice pleine, 110
  
- matrice tangente, 203
- matrice élémentaire, 86
- matériau linéaire élastique, 221
- matériau orthotrope, 225
- matériau semi-déformable, 279
- mesure, 11
- modèle de Saint-Venant-Kirchoff , 300
- modèle néo-hookéen, 300
- module de compressibilité, 303
- module de Young, 223
- méthode de Newton, 182
- méthode de continuation, 192
- méthode de Petrov-Galerkin, 64
- méthode de Rayleigh-Ritz, 64
- méthode de Ritz, 64
- méthode en espace-temps, 197
- méthode GMRES, 191
- méthodes de points fixes, 182
  
- noeuds d'interpolation, 79
- noeuds de calcul, 79, 139
- noeuds géométriques, 79, 138
- nombre de Peclet, 210
- nombre de Reynolds, 268
- norme, 28
- norme induite, 29
- normes équivalentes, 35
  
- ordre de convergence, 168, 171
  
- perte de charge, 264
- Petrov-Galerkin, 214
- poids d'intégration, 378
- points d'intégration, 151, 378
- Poiseuille, 258, 263
- polynôme  $P_1$  non conforme, 370
- polynômes d'Hermite, 124
- premier tenseur de Piola-Kirchoff, 297
- première variation, 187
- pression suiveuse, 306, 307, 311
- problème en grande déformation, 289
- problème mêlé, 53

- problème non linéaire, 181  
 produit dyadique, 353  
 produit scalaire, 28
- quadratures de Hammer, 381
- relèvement des conditions aux limites, 32, 37  
 relèvement des conditions aux limites essentielles, 51
- schéma d'Euler explicite, 196  
 schéma d'Euler implicite, 195, 196  
 schéma de Crank-Nicholson, 196  
 schéma de différence arrière d'ordre 2, 196  
 schéma de Gear, 204  
 simple couche, 21  
 solution manufacturée, 172  
 suite de Cauchy, 29  
 support, 98, 139  
 support d'une fonction, 8  
 système élémentaire, 86, 233
- tableau *adres*, 81, 140  
 tableau *connec*, 80, 139  
 tableau *coor*, 80, 139  
 tableau *numer*, 81, 140  
 tableau d'adressage, 81  
 tableau de numérotation, 81  
 tenseur d'élasticité, 221, 275  
 tenseur de Cauchy, 297  
 tenseur de Cauchy-Green, 291  
 tenseur de déformation, 220, 275, 291  
 tenseur de déformation de Green-Lagrange, 296  
 tenseur de Finger, 304  
 tenseur de taux de déformation, 244  
 tenseur des contraintes de Cauchy, 219  
 tenseur gradient de déformation, 220, 290  
 théorème de la divergence, 343  
 théorème de Lax-Milgram, 46, 48, 49, 52  
 trace au bord, 32  
 trace d'un tenseur, 355  
 trace normale, 37
- transformation géométrique, 87
- upwinding, 212
- variable primaire, 85  
 variable secondaire, 85, 122, 141  
 vecteur d'adressage, 81  
 vecteur de contrainte, 245  
 vecteur de contraintes normales, 222  
 vecteur des degrés de liberté élémentaires, 86  
 vecteur des réactions, 159  
 vecteur des sollicitations élémentaires, 87  
 vecteur global des degrés de liberté, 81, 107, 152  
 vecteur résidu, 183