

ÉQUATIONS
DIFFÉRENTIELLES
ET MÉTHODES NUMÉRIQUES
POUR INGÉNIEURS

par

Rémi VAILLANCOURT

Notes pour le cours MAT 2784 B

HIVER 2012

Département de mathématiques et de statistique
Université d'Ottawa
Ottawa, ON, Canada K1N 6N5

le 6 décembre 2011

VAILLANCOURT, Rémi, Département de mathématiques et de statistique,
Université d'Ottawa, Ottawa (Ontario), Canada K1N 6N5

courriel : remi@uottawa.ca

page d'accueil : www.site.uottawa.ca/~remi

Avec le concours du Conseil de recherches en sciences naturelles et en génie
du Canada.

Table des matières

| | |
|---|----------|
| Partie 1. Équations différentielles et transformation de Laplace | 1 |
| Chapitre 1. Équations différentielles du premier ordre | 3 |
| 1.1. Concepts fondamentaux | 3 |
| 1.2. Équations séparables | 5 |
| 1.3. Équations à coefficients homogènes | 7 |
| 1.4. Équations exactes | 9 |
| 1.5. Facteurs d'intégration | 16 |
| 1.6. Équations linéaires | 20 |
| 1.7. Familles de courbes orthogonales | 23 |
| 1.8. Champ des tangentes et solutions approchées | 25 |
| 1.9. Existence et unicité de la solution | 26 |
| Chapitre 2. Équations différentielles linéaires d'ordre deux | 31 |
| 2.1. Équations linéaires homogènes | 31 |
| 2.2. Équations homogènes à coefficients constants | 31 |
| 2.3. Base de l'espace solution | 32 |
| 2.4. Solutions indépendantes | 34 |
| 2.5. Modélisation en mécanique | 37 |
| 2.6. Équations d'Euler–Cauchy | 41 |
| Chapitre 3. Équations différentielles linéaires d'ordre quelconque | 47 |
| 3.1. Équations homogènes | 47 |
| 3.2. Équations linéaires homogènes à coefficients constants | 54 |
| 3.3. Équations linéaires non homogènes | 58 |
| 3.4. Méthode des coefficients indéterminés | 59 |
| 3.5. Solution particulière par variation des paramètres | 63 |
| 3.6. Systèmes asservis | 69 |
| Chapitre 4. Systèmes d'équations différentielles linéaires | 75 |
| 4.1. Introduction | 75 |
| 4.2. Théorème d'existence et d'unicité | 77 |
| 4.3. Système fondamental | 78 |
| 4.4. Systèmes linéaires à coefficients constants | 81 |
| 4.5. Systèmes linéaires non homogènes | 89 |
| Chapitre 5. Transformation de Laplace | 93 |
| 5.1. Définition | 93 |
| 5.2. Transformées de dérivées et d'intégrales | 97 |
| 5.3. Déplacements en s et en t | 101 |

| | | |
|--|--|------------|
| 5.4. | La fonction delta de Dirac | 109 |
| 5.5. | Dérivée et intégrale de la transformée | 111 |
| 5.6. | Équation différentielle de Laguerre | 114 |
| 5.7. | Convolution | 116 |
| 5.8. | Fractions simples | 118 |
| 5.9. | Transformées de fonctions périodiques | 119 |
| Chapitre 6. Solutions séries | | 123 |
| 6.1. | La méthode | 123 |
| 6.2. | Fondements de la méthode des séries de puissances | 124 |
| 6.3. | Équation et polynômes de Legendre | 130 |
| 6.4. | Orthogonalité des polynômes de Legendre | 133 |
| 6.5. | Série de Fourier–Legendre | 136 |
| 6.6. | Une application : la quadrature gaussienne | 139 |
| Partie 2. Méthodes numériques | | 145 |
| Chapitre 7. Résolution d'équations nonlinéaires par récurrence | | 147 |
| 7.1. | Calculs arithmétiques sur ordinateurs | 147 |
| 7.2. | Quelques théorème du calcul différentiel et intégral | 150 |
| 7.3. | Méthode de la dichotomie | 150 |
| 7.4. | Récurrence de point fixe | 154 |
| 7.5. | Méthodes de Newton, de la sécante et de la position fausse | 159 |
| 7.6. | Accélération de la convergence d'Aitken–Steffensen | 167 |
| 7.7. | La méthode d'Horner et la division synthétique | 169 |
| 7.8. | La méthode de Müller | 171 |
| Chapitre 8. Interpolation et extrapolation | | 175 |
| 8.1. | Polynômes d'interpolation de Lagrange | 175 |
| 8.2. | Interpolation newtonnienne aux différences divisées | 177 |
| 8.3. | Interpolation prograde de Gregory–Newton | 180 |
| 8.4. | Polynômes d'interpolation rétrograde de Gregory–Newton | 183 |
| 8.5. | Polynômes d'interpolation d'Hermite | 184 |
| 8.6. | Interpolation par splines du 3ème degré | 185 |
| Chapitre 9. Différentiation et intégration numériques | | 189 |
| 9.1. | Différentiation numérique | 189 |
| 9.2. | L'effect des erreurs d'arrondi et de méthode | 190 |
| 9.3. | Extrapolation de Richardson | 192 |
| 9.4. | Intégrations numériques élémentaires | 194 |
| 9.5. | La méthode des points milieux (composée) | 197 |
| 9.6. | La méthode des trapèzes (composée) | 199 |
| 9.7. | La méthode de Simpson (composée) | 201 |
| 9.8. | Intégration de Romberg pour la méthode des trapèzes | 203 |
| 9.9. | Quadratures adaptées | 204 |
| 9.10. | Quadratures gaussiennes | 206 |
| Chapitre 10. Résolution numérique d'équations différentielles | | 209 |
| 10.1. | Problèmes à valeur initiale | 209 |
| 10.2. | Méthodes explicites à un pas | 210 |

| | |
|---|------------|
| 10.3. Méthodes de Runge–Kutta explicites d’ordre 2, 3 et 4 | 213 |
| 10.4. Convergence des méthodes numériques | 221 |
| 10.5. Stabilité absolue des méthodes numériques | 222 |
| 10.6. Stabilité des méthodes de Runge–Kutta | 223 |
| 10.7. Paires de méthodes de Runge–Kutta | 225 |
| 10.8. Méthodes prédicteurs-correcteurs multipas | 231 |
| 10.9. Systèmes différentiels raides | 243 |
| Chapitre 11. Calcul matriciel | 253 |
| 11.1. Solution LU de $A\mathbf{x} = \mathbf{b}$ | 253 |
| 11.2. La décomposition de Cholesky | 261 |
| 11.3. Normes matricielles | 265 |
| 11.4. Méthodes de récurrence | 267 |
| 11.5. Systèmes surdéterminés | 269 |
| 11.6. Valeurs propres et vecteurs propres | 272 |
| 11.7. La décomposition QR | 276 |
| 11.8. L’algorithme QR | 277 |
| 11.9. La décomposition selon les valeurs singulières | 277 |
| Partie 3. Exercices suivis des solutions des exercices marqués d’un astérisque | 279 |
| Chapitre 12. Exercices pour équadifs et transf. de Laplace | 281 |
| Exercices pour le chapitre premier | 281 |
| Exercices pour le chapitre 2 | 283 |
| Exercices pour le chapitre 3 | 284 |
| Exercices pour le chapitre 4 | 286 |
| Exercices pour le chapitre 5 | 287 |
| Exercices pour le chapitre 6 | 289 |
| Chapitre 13. Exercices pour les méthodes numériques | 293 |
| Exercices pour le chapitre 7 | 293 |
| Exercices pour le chapitre 8 | 295 |
| Exercices pour le chapitre 9 | 296 |
| Exercices pour le chapitre 10 | 298 |
| Exercices pour le chapitre 11 | 299 |
| Chapitre 14. Solutions des exercices marqués d’un astérisque | 303 |
| Solutions des exercices du chapitre premier | 303 |
| Solutions des exercices du chapitre 2 | 305 |
| Solutions des exercices du chapitre 3 | 306 |
| Solutions des exercices du chapitre 4 | 308 |
| Solutions des exercices du chapitre 5 | 309 |
| Solutions des exercices du chapitre 6 | 311 |
| Solutions des exercices du chapitre 7 | 312 |
| Solutions des exercices du chapitre 8 | 314 |
| Solutions des exercices du chapitre 10 | 315 |
| Solutions des exercices du chapitre 11 | 321 |
| Chapitre 15. Formulaire et tables | 325 |

| | | |
|-------|--|-----|
| 15.1. | Facteur d'intégration de $M(x, y) dx + N(x, y) dy = 0$ | 325 |
| 15.2. | Les polynômes de LEGENDRE $P_n(x)$ sur $[-1, 1]$ | 325 |
| 15.3. | Les polynômes de LAGUERRE sur $0 \leq x < \infty$ | 326 |
| 15.4. | Développements de FOURIER-LEGENDRE | 327 |
| 15.5. | Table d'intégrales | 328 |
| 15.6. | La transformée de Laplace | 328 |
| | Index | 331 |

Partie 1

Équations différentielles et transformation de Laplace

Équations différentielles du premier ordre

1.1. Concepts fondamentaux

(a) Une *équation différentielle* est une relation entre une inconnue y et ses dérivées par rapport à une seule variable indépendante x (ou t).

Voici trois équations différentielles où $' := \frac{d}{dx}$:

- (1) $y' = \cos x$,
- (2) $y'' + 4y = 0$,
- (3) $x^2 y''' y' + 2 e^x y'' = (x^2 + 2)y^2$.

(b) L'ordre d'une équation différentielle est égal à l'ordre de la dérivée d'ordre le plus élevé.

Les équations (1), (2) et (3) ci-haut sont respectivement d'ordre 1, 2 et 3.

Voici une *équation aux dérivées partielles* du 2ème ordre :

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

(c) Une *solution explicite* d'une équation différentielle en x sur $]a, b[$ est une fonction d'une variable, $y = g(x)$, décrivant une courbe, telle que l'équation différentielle devient une identité en x sur $]a, b[$ quand on substitue $g(x)$, $g'(x)$, etc. à y , y' , etc. dans l'équation différentielle.

On voit que la fonction

$$y(x) = g(x) = e^{2x}$$

est une solution explicite de l'équation différentielle

$$\frac{dy}{dx} = 2y.$$

En effet, on a

$$\begin{aligned} \text{1er M} &:= y' = g'(x) = 2e^{2x}, \\ \text{2ème M} &:= 2y = 2g(x) = 2e^{2x}. \end{aligned}$$

Donc,

$$\text{1er M} = \text{2ème M}, \quad \text{pour tout } x.$$

On a donc une identité en x sur $] - \infty, \infty[$. □

(d) Une *solution implicite* d'une équation différentielle est une courbe définie par une équation de la forme $G(x, y) = c$ où c est une constante arbitraire.

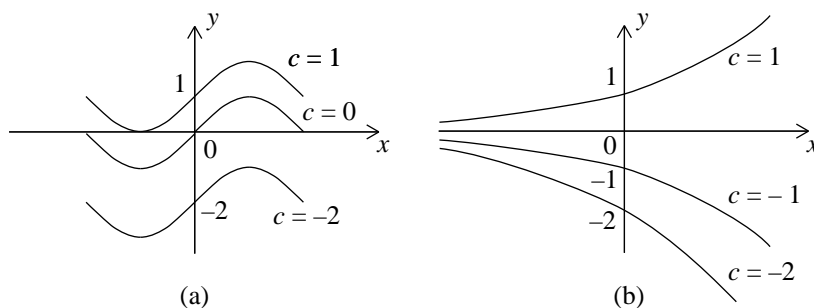


FIGURE 1.1. (a) Deux familles de courbes : (a) $y = \sin x + c$; (b) $y = c \exp(x)$.

On remarque qu'une solution implicite contient toujours le signe égal “=” suivi d'une constante, sinon $z = G(x, y)$ représente une surface et non une courbe.

On voit que la courbe

$$x^2 + y^2 - 1 = 0, \quad y > 0$$

dans le plan des xy est une solution implicite de l'équation différentielle

$$yy' = -x, \quad \text{sur } -1 < x < 1.$$

En effet, en considérant y comme fonction de x et en dérivant l'équation de la courbe par rapport à x ,

$$\frac{d}{dx}(x^2 + y^2 - 1) = \frac{d}{dx} 0 = 0,$$

on a

$$2x + 2yy' = 0 \quad \text{ou bien} \quad yy' = -x. \quad \square$$

(e) La *solution générale* d'une équation différentielle d'ordre n contient n constantes arbitraires.

La famille de fonctions

$$y(x) = \sin x + c$$

est la solution générale de l'équation différentielle du 1er ordre

$$y' = \cos x.$$

Pour $c = 1$ fixé, on obtient la solution unique,

$$y(x) = \sin x + 1,$$

qui passe par le point $(0, 1)$ de \mathbb{R}^2 . Etant donné un point quelconque (x_0, y_0) du plan, il y a une et une seule courbe de la famille donnée qui passe par ce point (V. figure 1.1(a)).

On voit de la même façon que la famille de fonctions

$$y(x) = c e^x$$

est la solution générale de l'équation différentielle

$$y' = y.$$

En fixant $c = -1$, on obtient la solution unique,

$$y(x) = -e^x,$$

qui passe par le point $(0, -1)$ de \mathbb{R}^2 . Etant donné un point quelconque (x_0, y_0) du plan, il y a une et une seule courbe de la famille donnée qui passe par ce point (V. figure 1.1(b)).

1.2. Équations séparables

Considérons une équation différentielle séparable de la forme

$$g(y) \frac{dy}{dx} = f(x). \quad (1.1)$$

On sépare l'équation réécrite sous forme de différentielles en mettant au 1er membre tous les termes en y et au second membre tous les termes en x :

$$g(y) dy = f(x) dx. \quad (1.2)$$

La solution d'une équation séparée s'obtient au moyen d'une intégrale indéfinie (primitive ou antiderivée) de chacun des membres à laquelle on ajoute une constante :

$$\int g(y) dy = \int f(x) dx + c, \quad (1.3)$$

c'est-à-dire

$$G(y) = F(x) + c, \quad \text{ou bien} \quad K(x, y) = -F(x) + G(y) = c.$$

Ces deux formes implicites de la solution définissent y comme fonction de x ou x comme fonction de y .

Supposons que $y = y(x)$ soit fonction de x et vérifions que (1.3) est bien une solution de (1.1) :

$$\begin{aligned} \frac{d}{dx}(\text{1er M}) &= \frac{d}{dx}G(y(x)) = G'(y(x))y'(x) = g(y)y', \\ \frac{d}{dx}(\text{2ème M}) &= \frac{d}{dx}[F(x) + c] = F'(x) = f(x). \quad \square \end{aligned}$$

EXEMPLE 1.1. Résoudre $y' = 1 + y^2$.

RÉSOLUTION. Puisque l'équation différentielle est séparable, on a

$$\int \frac{dy}{1+y^2} = \int dx + c \implies \arctan y = x + c.$$

Alors

$$y(x) = \tan(x + c)$$

est la solution générale, puisqu'elle contient une constante arbitraire. \square

EXEMPLE 1.2. Résoudre $y' = -2xy$ avec la condition initiale $y(0) = y_0$.

RÉSOLUTION. Puisque l'équation différentielle est séparable, la solution générale est

$$\int \frac{dy}{y} = - \int 2x dx + c_1 \implies \ln |y| = -x^2 + c_1.$$

En prenant l'exponentielle de la solution, on a

$$y(x) = e^{-x^2+c_1} = e^{c_1} e^{-x^2}$$

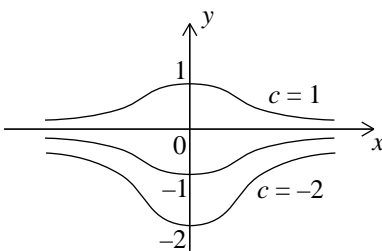


FIGURE 1.2. Trois fonctions cloches.

qu'on récrit sous la forme

$$y(x) = c e^{-x^2}.$$

On remarquera que la constante additive c_1 est devenue une constante multiplicative après l'exponentiation. Dans la figure 1.2 on voit trois fonctions cloches membres de la famille de la solution générale.

Enfin, la solution qui satisfait la condition initiale, est

$$y(x) = y_0 e^{-x^2}.$$

Cette solution est unique. □

EXEMPLE 1.3. D'après la loi du refroidissement de Newton, le taux de changement de la température $T(t)$ d'un corps dans un milieu environnant de température T_0 est proportionnel à la différence $T(t) - T_0$ des températures :

$$\frac{dT}{dt} = -k(T - T_0).$$

On plonge une boule de cuivre dans un grand bassin de liquide dont la température est maintenue à 30 degrés. Si la température initiale de la boule est de 100 degrés et si sa température après 3 min est de 70 degrés, quand sera-t-elle de 31 degrés?

RÉSOLUTION. On a l'équation différentielle séparable :

$$\frac{dT}{dt} = -k(T - 30) \implies \frac{dT}{T - 30} = -k dt.$$

Alors

$$\ln |T - 30| = -kt + c_1 \quad (\text{constante additive})$$

$$T - 30 = e^{c_1 - kt} = c e^{-kt} \quad (\text{constante multiplicative})$$

$$T(t) = 30 + c e^{-kt}.$$

À $t = 0$,

$$100 = 30 + c \implies c = 70.$$

À $t = 3$,

$$70 = 30 + 70 e^{-3k} \implies e^{-3k} = \frac{4}{7}.$$

Si $T(t) = 31$, alors

$$31 = 70 (e^{-3k})^{t/3} + 30 \implies (e^{-3k})^{t/3} = \frac{1}{70}.$$

En prenant le logarithme des deux membres, on obtient

$$\frac{t}{3} \ln\left(\frac{4}{7}\right) = \ln\left(\frac{1}{70}\right),$$

d'où il vient que

$$t = 3 \frac{\ln(1/70)}{\ln(4/7)} = 3 \times \frac{-4.25}{-0.56} = 22.78 \text{ min} \quad \square$$

1.3. Équations à coefficients homogènes

DÉFINITION 1.1. Une fonction $M(x, y)$ est *homogène de degré s simultanément en x et en y* si

$$M(\lambda x, \lambda y) = \lambda^s M(x, y), \quad \text{pour tout } x, y, \lambda. \quad (1.4)$$

Les équations différentielles à coefficients homogènes du même degré sont séparables de la façon suivante.

THÉORÈME 1.1. *Soit une équation différentielle à coefficients homogènes de degré s ,*

$$M(x, y)dx + N(x, y)dy = 0. \quad (1.5)$$

Alors chacune des substitutions $y = xu$ et $x = yu$ rend l'équation différentielle séparable.

DÉMONSTRATION. Posons

$$y = xu, \quad dy = x du + u dx,$$

dans (1.5). Alors,

$$\begin{aligned} M(x, xu) dx + N(x, xu)[x du + u dx] &= 0, \\ x^s M(1, u) dx + x^s N(1, u)[x du + u dx] &= 0, \\ [M(1, u) + uN(1, u)] dx + xN(1, u) du &= 0. \end{aligned}$$

On sépare cette dernière équation :

$$\frac{N(1, u)}{M(1, u) + uN(1, u)} du = -\frac{dx}{x},$$

qui admet la solution générale

$$\int \frac{N(1, u)}{M(1, u) + uN(1, u)} du = -\ln|x| + c. \quad \square$$

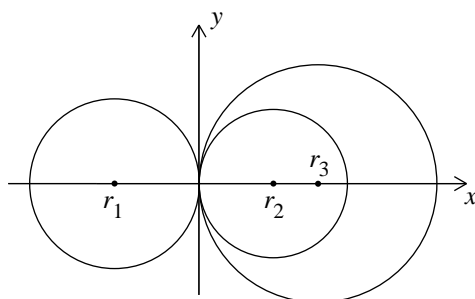
EXEMPLE 1.4. Résoudre $2xyy' - y^2 + x^2 = 0$.

RÉSOLUTION. On récrit l'équation sous forme différentielle :

$$(x^2 - y^2) dx + 2xy dy = 0.$$

Puisque les coefficients sont des fonctions homogènes de degré 2 en x et y , posons

$$x = yu, \quad dx = y du + u dy.$$

FIGURE 1.3. Familles de cercles de centre $(r, 0)$.

Alors la dernière équation devient

$$\begin{aligned}(y^2 u^2 - y^2)[y du + u dy] + 2y^2 u dy &= 0, \\ (u^2 - 1)[y du + u dy] + 2u dy &= 0, \\ (u^2 - 1)y du + [(u^2 - 1)u + 2u] dy &= 0, \\ \frac{u^2 - 1}{u(u^2 + 1)} du &= -\frac{dy}{y}.\end{aligned}$$

Puisque le 1er membre de cette équation semble difficile à intégrer, recommençons avec la substitution

$$y = xu, \quad dy = x du + u dx.$$

Alors,

$$\begin{aligned}(x^2 - x^2 u^2) dx + 2x^2 u[x du + u dx] &= 0, \\ [(1 - u^2) + 2u^2] dx + 2ux du &= 0, \\ \int \frac{2u}{1 + u^2} du &= -\int \frac{dx}{x} + c_1.\end{aligned}$$

Cette dernière équation s'intègre facilement :

$$\begin{aligned}\ln(u^2 + 1) &= -\ln|x| + c_1, \\ \ln|x(u^2 + 1)| &= c_1, \\ x \left[\left(\frac{y}{x}\right)^2 + 1 \right] &= e^{c_1} = c.\end{aligned}$$

Alors la solution générale est

$$y^2 + x^2 = cx.$$

Si l'on pose $c = 2r$ dans la formule précédente et l'on ajoute r^2 aux deux membres, on obtient

$$(x - r)^2 + y^2 = r^2.$$

La solution générale décrit donc une famille de cercles de centre $(r, 0)$ et de rayon $|r|$ (V. figure 1.3). \square

EXEMPLE 1.5. Résoudre l'équation différentielle

$$y' = g\left(\frac{y}{x}\right).$$

RÉSOLUTION. Récrivons cette équation sous forme différentielle :

$$g\left(\frac{y}{x}\right) dx - dy = 0.$$

C'est une équation aux coefficients homogènes de degré zéro où la fonction g est homogène de degré zéro en x et y . Si l'on pose

$$y = xu, \quad dy = x du + u dx,$$

la dernière équation devient

$$\begin{aligned} g(u) dx - x du - u dx &= 0, \\ x du &= [g(u) - u] dx, \\ \frac{du}{g(u) - u} &= \frac{dx}{x}, \end{aligned}$$

qui est une équation séparée. On peut donc l'intégrer directement :

$$\int \frac{du}{g(u) - u} = \int \frac{dx}{x} + c,$$

et substituer $u = y/x$ dans la solution après intégration. \square

1.4. Équations exactes

DÉFINITION 1.2. L'équation différentielle du 1er ordre

$$M(x, y) dx + N(x, y) dy = 0 \tag{1.6}$$

est *exacte* si le 1er membre est une différentielle totale ou exacte d'une fonction $u(x, y)$:

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy. \tag{1.7}$$

Si (1.6) est exacte, alors

$$du = 0$$

et la solution générale de (1.6) est

$$u(x, y) = c. \tag{1.8}$$

Si l'on compare les expressions (1.6) et (1.7), on voit que

$$\frac{\partial u}{\partial x} = M, \quad \frac{\partial u}{\partial y} = N. \tag{1.9}$$

Le théorème suivant donne une condition nécessaire et suffisante pour que l'équation (1.6) soit exacte.

THÉORÈME 1.2. *Soit $M(x, y)$ et $N(x, y)$ deux fonctions continûment dérivables sur un ensemble $\Omega \in \mathbb{R}^2$ connexe et simplement connexe (c'est-à-dire d'un seul morceau et sans trou). Alors l'équation différentielle*

$$M(x, y) dx + N(x, y) dy = 0 \tag{1.10}$$

est exacte si et seulement si

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}, \quad \text{sur } \Omega. \tag{1.11}$$

DÉMONSTRATION. **Nécessité:** Supposons que (1.10) est exacte; alors

$$\frac{\partial u}{\partial x} = M, \quad \frac{\partial u}{\partial y} = N.$$

Par conséquent,

$$\frac{\partial M}{\partial y} = \frac{\partial^2 u}{\partial y \partial x} = \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial N}{\partial x},$$

où l'échange de l'ordre des dérivées en x et y est justifiée par la continuité des deux termes extrêmes.

Suffisance: On suppose (1.11) et l'on construit une fonction $F(x, y)$ telle que

$$dF(x, y) = M(x, y) dx + N(x, y) dy.$$

Soit $\varphi(x, y) \in C^2(\Omega)$ une fonction telle que

$$\frac{\partial \varphi}{\partial x} = M.$$

On peut prendre, par exemple,

$$\varphi(x, y) = \int M(x, y) dx, \quad y \text{ fixé.}$$

Alors,

$$\begin{aligned} \frac{\partial^2 \varphi}{\partial y \partial x} &= \frac{\partial M}{\partial y} \\ &= \frac{\partial N}{\partial x}, \quad \text{par (1.11).} \end{aligned}$$

Puisque

$$\frac{\partial^2 \varphi}{\partial y \partial x} = \frac{\partial^2 \varphi}{\partial x \partial y}$$

par la continuité des deux membres, il suit que

$$\frac{\partial^2 \varphi}{\partial x \partial y} = \frac{\partial N}{\partial x}.$$

Si l'on intègre en x , on obtient

$$\begin{aligned} \frac{\partial \varphi}{\partial y} &= \int \frac{\partial^2 \varphi}{\partial x \partial y} dx = \int \frac{\partial N}{\partial x} dx, \quad y \text{ fixé,} \\ &= N(x, y) + B'(y). \end{aligned}$$

Prenons

$$F(x, y) = \varphi(x, y) - B(y).$$

Alors

$$\begin{aligned} dF &= \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy - B'(y) dy \\ &= M dx + N dy + B'(y) dy - B'(y) dy \\ &= M dx + N dy. \quad \square \end{aligned}$$

On illustre par des exemples une **méthode pratique** pour résoudre des équations exactes.

EXEMPLE 1.6. Trouver la solution générale de l'équation différentielle

$$3x(xy - 2) dx + (x^3 + 2y) dy = 0$$

et la solution qui satisfait la condition initiale $y(1) = -1$. Tracer la solution sur $1 \leq x \leq 4$.

RÉSOLUTION. (a) **Résolution analytique par la méthode pratique.**—
On vérifie que l'équation est exacte :

$$\begin{aligned} M &= 3x^2y - 6x, & N &= x^3 + 2y, \\ \frac{\partial M}{\partial y} &= 3x^2, & \frac{\partial N}{\partial x} &= 3x^2, \\ \frac{\partial M}{\partial y} &= \frac{\partial N}{\partial x}. \end{aligned}$$

Il suit que l'équation est exacte. On peut donc l'intégrer. De

$$\frac{\partial u}{\partial x} = M,$$

on a

$$\begin{aligned} u(x, y) &= \int M(x, y) dx + T(y), & y \text{ fixé,} \\ &= \int (3x^2y - 6x) dx + T(y) \\ &= x^3y - 3x^2 + T(y). \end{aligned}$$

De

$$\frac{\partial u}{\partial y} = N,$$

on a

$$\begin{aligned} \frac{\partial u}{\partial y} &= \frac{\partial}{\partial y}(x^3y - 3x^2 + T(y)) \\ &= x^3 + T'(y) = N \\ &= x^3 + 2y. \end{aligned}$$

Donc

$$T'(y) = 2y.$$

Il est essentiel que $T'(y)$ soit fonction de y seulement, sinon il y a erreur quelque part : ou l'équation n'est pas exacte ou on a une erreur de calcul.

On intègre $T'(y)$:

$$T(y) = y^2.$$

Il n'est pas nécessaire d'ajouter une constante d'intégration ici parce qu'on aura une constante dans $u(x, y) = c$. On a donc la **surface**

$$u(x, y) = x^3y - 3x^2 + y^2.$$

Puisque $du = 0$, alors $u(x, y) = c$ et la solution générale implicite, qui contient *une constante arbitraire et un signe " = "*, c'est-à-dire une **courbe**, est donc

$$x^3y - 3x^2 + y^2 = c.$$

On détermine la constante c au moyen de la condition initiale $y(1) = -1$. On pose donc $x = 1$ et $y = -1$ dans la solution générale et l'on obtient

$$c = -3.$$

Alors, la solution implicite qui satisfait la condition initiale est

$$x^3y - 3x^2 + y^2 = -3.$$

(b) Résolution par Matlab symbolique.— La solution générale est :

```
>> y = dsolve('(x^3+2*y)*Dy=-3*x*(x*y-2)', 'x')
y =
[ -1/2*x^3+1/2*(x^6+12*x^2+4*C1)^(1/2)]
[ -1/2*x^3-1/2*(x^6+12*x^2+4*C1)^(1/2)]
```

La solution du problème à valeur initiale est la branche inférieure avec $C1 = -3$, comme on peut voir en insérant la condition initiale ' $y(1)=-1$ ' dans la commande précédente :

```
>> y = dsolve('(x^3+2*y)*Dy=-3*x*(x*y-2)', 'y(1)=-1', 'x')
y = -1/2*x^3-1/2*(x^6+12*x^2-12)^(1/2)
```

(c) Solution du problème à valeur initiale par Matlab numérique.— On emploie la condition initiale $y(1) = -1$. Le fichier M `exp1_6.m` est

```
function yprime = exp1_6(x,y); %MAT 2784, Exp 1.6.
yprime = -3*x*(x*y-2)/(x^3+2*y);
```

L'appel du solveur `ode23` et de la commande `plot` sont :

```
>> xspan = [1 4]; % solution pour 1<=x<=4
>> y0 = -1; % condition initiale
>> [x,y] = ode23(@exp1_6,xspan,y0);%xspan en Matlab 7.4
>> subplot(2,2,1); plot(x,y);
>> title('Graphe de la solution pour l''exemple 1.6');
>> xlabel('x'); ylabel('y(x)');
>> print -deps2 Fig.exp1.6
```

□

EXEMPLE 1.7. Trouver la solution générale de l'équation différentielle

$$(2x^3 - xy^2 - 2y + 3) dx - (x^2y + 2x) dy = 0$$

et la solution qui satisfait la condition initiale $y(1) = -1$. Tracer la solution sur $1 \leq x \leq 4$.

RÉSOLUTION. **(a) Résolution analytique par la méthode pratique.**— Remarquer que $N(x, y) = -(x^2y + 2x)$ du fait que le 1er membre de l'équation est $M dx + N dy$. On vérifie que l'équation est exacte :

$$\frac{\partial M}{\partial y} = -2xy - 2, \quad \frac{\partial N}{\partial x} = -2xy - 2,$$

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}.$$

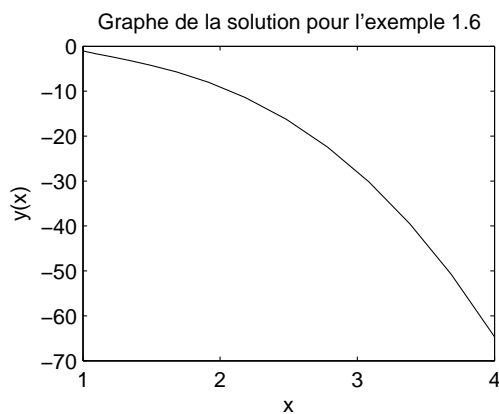


FIGURE 1.4. Graphe de la solution pour l'exemple 1.6.

Il suit que l'équation est exacte. On peut donc l'intégrer. De

$$\frac{\partial u}{\partial y} = N,$$

on a

$$\begin{aligned} u(x, y) &= \int N(x, y) dy + T(x), \quad x \text{ fixé,} \\ &= \int (-x^2 y - 2x) dy + T(x) \\ &= -\frac{x^2 y^2}{2} - 2xy + T(x). \end{aligned}$$

De

$$\frac{\partial u}{\partial x} = M,$$

on a

$$\begin{aligned} \frac{\partial u}{\partial x} &= -xy^2 - 2y + T'(x) = M \\ &= 2x^3 - xy^2 - 2y + 3. \end{aligned}$$

Donc

$$T'(x) = 2x^3 + 3.$$

Il est essentiel que $T'(x)$ soit fonction de x seulement, sinon il y a erreur quelque part : ou l'équation n'est pas exacte ou on a une erreur de calcul.

On intègre $T'(x)$:

$$T(x) = \frac{x^4}{2} + 3x.$$

Il n'est pas nécessaire d'ajouter une constante d'intégration ici parce qu'on aura une constante dans $u(x, y) = c$. On a donc la **surface**

$$u(x, y) = -\frac{x^2 y^2}{2} - 2xy + \frac{x^4}{2} + 3x.$$

Puisque $du = 0$, alors $u(x, y) = c$ et la solution générale implicite, qui contient une constante arbitraire et un signe "=", c'est-à-dire une **courbe**, est donc

$$x^4 - x^2y^2 - 4xy + 6x = c.$$

Puisque $y = -1$ en $x = 1$, on obtient

$$c = 10.$$

Alors, la solution implicite qui satisfait la condition initiale est

$$x^4 - x^2y^2 - 4xy + 6x = 10.$$

(b) Résolution par Matlab symbolique.— La solution générale est :

```
>> y = dsolve('(x^2*y+2*x)*Dy=(2*x^3-x*y^2-2*y+3)', 'x')
y =
[ (-2-(4+6*x+x^4+2*C1)^(1/2))/x]
[ (-2+(4+6*x+x^4+2*C1)^(1/2))/x]
```

La solution du problème à valeur initiale est la branche inférieure avec $C1 = -5$:

```
>> y = dsolve('(x^2*y+2*x)*Dy=(2*x^3-x*y^2-2*y+3)', 'y(1)=-1', 'x')
y =(-2+(-6+6*x+x^4)^(1/2))/x
```

(c) Solution du problème à valeur initiale par Matlab numérique.— On emploie la condition initiale $y(1) = -1$. Le fichier M `exp1_7.m` est

```
function yprime = exp1_7(x,y); %MAT 2784, Exp 1.7.
yprime = (2*x^3-x*y^2-2*y+3)/(x^2*y+2*x);
```

L'appel du solveur `ode23` et de la commande `plot` sont :

```
>> xspan = [1 4]; % solution pour 1<=x<=4
>> y0 = -1; % condition initiale
>> [x,y] = ode23(@exp1_7,xspan,y0);%xspan en Matlab 7.4
>> subplot(2,2,1); plot(x,y);
>> title('Graphe de la solution pour l''exemple 1.7');
>> xlabel('x'); ylabel('y(x)');
>> print -deps2 Fig.exp1.7
```

□

Dans la suite, on notera les dérivées partielles :

$$u_x(x, y) := \frac{\partial u}{\partial x}, \quad u_y(x, y) := \frac{\partial u}{\partial y}.$$

L'exemple suivant montre que la méthode flanche si l'on tente d'intégrer une équation qui n'est pas exacte.

EXEMPLE 1.8. Résoudre

$$x dy - y dx = 0.$$

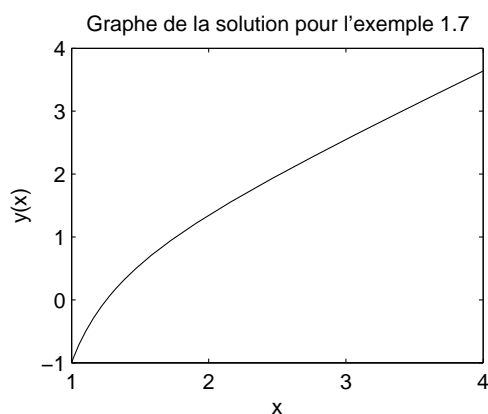


FIGURE 1.5. Graphe de la solution pour l'exemple 1.7.

RÉSOLUTION. On récrit l'équation sous forme standard :

$$y dx - x dy = 0.$$

L'équation n'est pas exacte parce que

$$M_y = 1 \neq -1 = N_x.$$

Essayons de résoudre l'équation par la méthode proposée :

$$\begin{aligned} u(x, y) &= \int u_x(x, y) dx = \int M dx = \int y dx = yx + T(y), \\ u_y(x, y) &= x + T'(y) = N = -x. \end{aligned}$$

Donc

$$T'(y) = -2x.$$

Ceci est impossible parce que $T(y)$ doit être fonction de y seulement. □

EXEMPLE 1.9. Soit l'équation différentielle

$$(ax + by) dx + (kx + ly) dy = 0.$$

Choisir a, b, k, l pour que l'équation soit exacte.

RÉSOLUTION.

$$M_y = b, \quad N_x = k \implies k = b.$$

$$\begin{aligned} u(x, y) &= \int u_x dx = \int M dx = \int (ax + by) dx = \frac{ax^2}{2} + bxy + T(y), \\ u_y(x, y) &= bx + T'(y) = N = bx + ly \implies T'(y) = ly \implies T(y) = \frac{ly^2}{2}. \end{aligned}$$

Donc

$$u(x, y) = \frac{ax^2}{2} + bxy + \frac{ly^2}{2}, \quad a, b, l \text{ arbitraires.}$$

La solution générale est

$$\frac{ax^2}{2} + bxy + \frac{ly^2}{2} = c_1 \quad \text{ou bien} \quad ax^2 + 2bxy + ly^2 = c. \quad \square$$

1.5. Facteurs d'intégration

Si l'équation différentielle

$$M(x, y) dx + N(x, y) dy = 0 \quad (1.12)$$

n'est pas exacte, on peut la rendre exacte en la multipliant par un facteur d'intégration $\mu(x, y)$:

$$\mu(x, y)M(x, y) dx + \mu(x, y)N(x, y) dy = 0. \quad (1.13)$$

L'équation (1.13) sera exacte si

$$\mu_y M + \mu M_y = \mu_x N + \mu N_x. \quad (1.14)$$

En général, il est difficile de résoudre l'équation aux dérivées partielles (1.14).

On considère les deux cas particuliers où μ est une fonction d'une seule variable : $\mu = \mu(x)$ ou $\mu = \mu(y)$.

Cas 1. Si $\mu = \mu(x)$ est une fonction de x seulement, alors $\mu_x = \mu'(x)$ et $\mu_y = 0$. Donc (1.14) devient une équation différentielle :

$$N\mu'(x) = \mu(M_y - N_x). \quad (1.15)$$

Si le 1er membre de l'expression suivante :

$$\frac{M_y - N_x}{N} = f(x) \quad (1.16)$$

est une fonction de x seulement, alors (1.15) est séparable :

$$\frac{d\mu}{\mu} = \frac{M_y - N_x}{N} dx = f(x) dx.$$

En intégrant cette équation séparée, on obtient le facteur d'intégration

$$\mu(x) = e^{\int f(x) dx}. \quad (1.17)$$

Cas 2. De même, si $\mu = \mu(y)$ est une fonction de y seulement, alors $\mu_x = 0$ et $\mu_y = \mu'(y)$. Donc (1.14) devient une équation différentielle :

$$M\mu'(y) = -\mu(M_y - N_x). \quad (1.18)$$

Si le 1er membre de l'expression suivante

$$\frac{M_y - N_x}{M} = g(y) \quad (1.19)$$

est une fonction de y seulement, alors (1.18) est séparable :

$$\frac{d\mu}{\mu} = -\frac{M_y - N_x}{M} dy = -g(y) dy.$$

En intégrant cette équation séparée, on obtient le facteur d'intégration

$$\mu(y) = e^{-\int g(y) dy}. \quad (1.20)$$

Remarquer la présence du signe moins dans (1.20) et son absence dans (1.17).

EXEMPLE 1.10. Résoudre

$$(4xy + 3y^2 - x) dx + x(x + 2y) dy = 0.$$

RÉSOLUTION. (a) **Résolution analytique.**— Cette équation n'est pas exacte puisque

$$M_y = 4x + 6y \neq N_x = 2x + 2y.$$

Puisque

$$\frac{M_y - N_x}{N} = \frac{2x + 4y}{x(x + 2y)} = \frac{2(x + 2y)}{x(x + 2y)} = \frac{2}{x} = f(x)$$

est une fonction de x seulement, on a le facteur d'intégration

$$\mu(x) = e^{\int \frac{2}{x} dx} = e^{2 \ln x} = e^{\ln x^2} = x^2.$$

Multipliant l'équation différentielle par x^2 , on obtient l'équation exacte

$$\mu M dx + \mu N dy = x^2(4xy + 3y^2 - x) dx + x^3(x + 2y) dy = 0.$$

On résout cette équation par la méthode pratique :

$$\begin{aligned} u(x, y) &= \int (x^4 + 2x^3y) dy + T(x) \\ &= x^4y + x^3y^2 + T(x), \\ u_x(x, y) &= 4x^3y + 3x^2y^2 + T'(x) = \mu M \\ &= 4x^3y + 3x^2y^2 - x^3. \end{aligned}$$

Alors

$$T'(x) = -x^3 \implies T(x) = -\frac{x^4}{4}.$$

Pas de constante d'intégration ici; elle apparaîtra plus loin. On a donc

$$u(x, y) = x^4y + x^3y^2 - \frac{x^4}{4}$$

et la solution générale est

$$x^4y + x^3y^2 - \frac{x^4}{4} = c_1 \quad \text{ou bien} \quad 4x^4y + 4x^3y^2 - x^4 = c.$$

(b) **Résolution par Matlab symbolique.**— Matlab ne trouve pas la solution générale de l'équation inexacte :

```
>> y = dsolve('x*(x+2*y)*Dy=-(4*x+3*y^2-x)', 'x')
Warning: Explicit solution could not be found.
> In HD2:Matlab5.1:Toolbox:symbolic:dsolve.m at line 200
y = [ empty sym ]
```

mais résout l'équation exacte :

```
>> y = dsolve('x^2*(x^3+2*y)*Dy=-3*x^3*(x*y-2)', 'x')
y =
[ -1/2*x^3-1/2*(x^6+12*x^2+4*C1)^(1/2) ]
[ -1/2*x^3+1/2*(x^6+12*x^2+4*C1)^(1/2) ]
```

□

EXEMPLE 1.11. Résoudre

$$y(x + y + 1) dx + x(x + 3y + 2) dy = 0.$$

RÉSOLUTION. (a) **Résolution analytique.**— Puisque

$$\frac{M_y - N_x}{N} = \frac{-x - y - 1}{x(x + 3y + 2)} \neq f(x),$$

on essaie

$$\frac{M_y - N_x}{M} = \frac{-(x + y + 1)}{y(x + y + 1)} = -\frac{1}{y} = g(y),$$

qui est fonction de y seulement. On a le facteur d'intégration

$$\mu(y) = e^{-\int g(y) dy} = e^{\int \frac{1}{y} dy} = e^{\ln y} = y.$$

Multipliant l'équation différentielle par y , on obtient l'équation exacte

$$\mu M dx + \mu N dy = (xy^2 + y^3 + y^2) dx + (x^2y + 3xy^2 + 2xy) dy = 0.$$

On résout cette équation par la méthode pratique :

$$\begin{aligned} u(x, y) &= \int (xy^2 + y^3 + y^2) dx + T(y) \\ &= \frac{x^2y^2}{2} + xy^3 + xy^2 + T(y), \\ u_y(x, y) &= x^2y + 3xy^2 + 2xy + T'(y) = \mu N \\ &= x^2y + 3xy^2 + 2xy. \end{aligned}$$

Alors

$$T'(y) = 0 \implies T(y) = k = 0.$$

On a donc

$$u(x, y) = \frac{x^2y^2}{2} + xy^3 + xy^2$$

et la solution générale est

$$\frac{x^2y^2}{2} + xy^3 + xy^2 = c_1 \quad \text{ou bien} \quad x^2y^2 + 2xy^3 + 2xy^2 = c.$$

(b) **Résolution par Matlab symbolique.**— La commande `dsolve` de Matlab symbolique produit une solution générale très complexe pour l'équation inexacte et pour l'équation exacte. Ces solutions ne se simplifient pas au moyen des commandes `simplify` et `simple`.

On reprend donc la méthode pratique et demande à Matlab symbolique de faire les simples manipulations algébriques et analytiques.

```
>> clear
>> syms M N x y u
>> M = y*(x+y+1); N = x*(x+3*y+2);
>> test = diff(M,'y') - diff(N,'x') % equation exacte ou non
    test = -x-y-1 % equation non exacte
>> syms mu g
>> g = (diff(M,'y') - diff(N,'x'))/M
    g = (-x-y-1)/y/(x+y+1)
>> g = simple(g)
    g = -1/y % une fonction de y seulement
>> mu = exp(-int(g,'y')) % facteur d'int\egration
    mu = y
>> syms MM NN
```



```

>> MM = mu*M; NN = mu*N; % multiplication par le facteur d'integration
>> u = int(MM, 'x') % solution u; T(y) arbitraire pas inclus
    u = y^2*(1/2*x^2+y*x+x)
>> syms DT
>> DT = simple(diff(u, 'y') - NN)
    DT = 0 % T'(y) = 0 implique T(y) = 0.
>> u = u
    u = y^2*(1/2*x^2+y*x+x) % solution generale u = c.

```

La solution générale est

$$\frac{x^2y^2}{2} + xy^3 + xy^2 = c_1 \quad \text{ou} \quad x^2y^2 + 2xy^3 + 2xy^2 = c. \quad \square$$

REMARQUE 1.1. Une équation séparée,

$$f(x) dx + g(y) dy = 0,$$

est exacte. En effet, $M_y = 0$ et $N_x = 0$, ce qui donne le facteur d'intégration

$$\mu(x) = e^{\int 0 dx} = 1, \quad \mu(y) = e^{-\int 0 dy} = 1.$$

Si l'on résout cette équation par la méthode pratique des équations exactes, on a :

$$\begin{aligned}
 u(x, y) &= \int f(x) dx + T(y), \\
 u_y(x, y) &= T'(y) = g(y) \implies T(y) = \int g(y) dy, \\
 u(x, y) &= \int f(x) dx + \int g(y) dy = c.
 \end{aligned}$$

On obtient donc la même solution que celle obtenue par la méthode (1.3).

REMARQUE 1.2. Le facteur qui transforme une équation séparable en une équation séparée est un facteur d'intégration puisque cette dernière est exacte.

EXEMPLE 1.12. Soit l'équation séparable

$$y' = 1 + y^2, \quad \text{c'est-à-dire} \quad (1 + y^2) dx - dy = 0.$$

Montrer que le facteur $(1 + y^2)^{-1}$ qui sépare l'équation est un facteur d'intégration.

RÉSOLUTION. On a

$$M_y = 2y, \quad N_x = 0, \quad \frac{2y - 0}{1 + y^2} = g(y).$$

Donc

$$\begin{aligned}
 \mu(y) &= e^{-\int \frac{2y}{1+y^2} dy} \\
 &= e^{\ln[(1+y^2)^{-1}]} = \frac{1}{1+y^2}. \quad \square
 \end{aligned}$$

Dans le prochain exemple, on trouve facilement un facteur d'intégration $\mu(x, y)$ qui est une fonction de x et de y .

EXEMPLE 1.13. Soit l'équation séparable

$$y dx + x dy = 0.$$

Montrer que le facteur

$$\mu(x, y) = \frac{1}{xy},$$

qui sépare l'équation, est un facteur d'intégration.

RÉSOLUTION. L'équation différentielle

$$\mu(x, y)y dx + \mu(x, y)x dy = \frac{1}{x} dx + \frac{1}{y} dy = 0$$

est séparée; donc elle est exacte. \square

1.6. Équations linéaires

Considérons l'équation non homogène du 1er ordre de la forme

$$y' + f(x)y = r(x). \quad (1.21)$$

Le 1er membre de cette équation est une expression linéaire en y et y' . On dira donc que (1.21) est une équation différentielle *linéaire*.

Dans la présente section, on résout (1.21) en transformant le 1er membre en une dérivée totale au moyen d'un facteur d'intégration. À l'exemple 3.10, on exprimera la solution générale de (1.21) comme la somme de la solution de l'équation homogène (avec second membre nul) et d'une solution particulière de l'équation non homogène. On considérera les solutions en série de puissances au chapitre 5 et les solutions numériques au chapitre 12

Suivant la première méthode, on multiplie (1.21) par $\mu(x)$,

$$\mu(x)(y' + f(x)y) = \mu(x)r(x),$$

pour transformer le 1er membre en une dérivée exacte ou totale de la forme

$$\begin{aligned} \frac{d}{dx}[\mu(x)y(x)] &= \mu'(x)y(x) + \mu(x)y'(x) \\ &= \mu(x)[y' + f(x)y]. \end{aligned}$$

Après simplification, on obtient l'équation séparable

$$\mu' = \mu f \quad \implies \quad \frac{d\mu}{\mu} = f(x) dx \quad \implies \quad \ln \mu(x) = \int f(x) dx.$$

On a donc le facteur d'intégration

$$\mu(x) = e^{\int f(x) dx}.$$

Donc l'équation sous forme différentielle devient

$$d \left[e^{\int f(x) dx} y(x) \right] = e^{\int f(x) dx} r(x) dx.$$

On intègre les deux membres par rapport à x :

$$e^{\int f(x) dx} y(x) = \int e^{\int f(x) dx} r(x) dx + c.$$

et l'on résout pour $y(x)$, ce qui donne la solution générale de (1.21) :

$$y(x) = e^{-\int f(x) dx} \left[\int e^{\int f(x) dx} r(x) dx + c \right]. \quad (1.22)$$

EXEMPLE 1.14. Résoudre l'équation différentielle linéaire

$$x^2 y' + 2xy = \sinh 3x.$$

RÉSOLUTION. On récrit l'équation sous forme standard :

$$y' + \frac{2}{x} y = \frac{1}{x^2} \sinh 3x.$$

Le facteur d'intégration est

$$\mu(x) = e^{\int \frac{2}{x} dx} = e^{\ln x^2} = x^2.$$

Alors,

$$d(x^2 y) = \sinh 3x dx, \quad \text{c'est-à-dire} \quad \frac{d}{dx}(x^2 y) = \sinh 3x.$$

On a donc

$$x^2 y(x) = \int \sinh 3x dx + c = \frac{1}{3} \cosh 3x + c,$$

ou

$$y(x) = \frac{1}{3x^2} \cosh 3x + \frac{c}{x^2}. \quad \square$$

EXEMPLE 1.15. Résoudre l'équation différentielle linéaire

$$y dx + (3x - xy + 2) dy = 0.$$

RÉSOLUTION. On récrit cette équation sous la forme d'une équation linéaire en $x(y)$:

$$\frac{dx}{dy} + \left(\frac{3}{y} - 1 \right) x = -\frac{2}{y}, \quad y \neq 0.$$

Le facteur d'intégration qui rend le 1er membre exact, est

$$\mu(y) = e^{\int (\frac{3}{y} - 1) dy} = e^{\ln y^3 - y} = y^3 e^{-y}.$$

On a donc

$$\frac{d}{dy}(y^3 e^{-y} x) = -2y^2 e^{-y}, \quad \text{c'est-à-dire} \quad d(y^3 e^{-y} x) = -2y^2 e^{-y} dy.$$

Alors

$$\begin{aligned} y^3 e^{-y} x &= -2 \int y^2 e^{-y} dy + c \\ &= 2y^2 e^{-y} - 4 \int y e^{-y} dy + c \\ &= 2y^2 e^{-y} + 4y e^{-y} - 4 \int e^{-y} dy + c \\ &= 2y^2 e^{-y} + 4y e^{-y} + 4e^{-y} + c. \end{aligned}$$

La solution générale est donc

$$xy^3 = 2y^2 + 4y + 4 + c e^y. \quad \square$$

On peut rendre linéaires certaines équations différentielles nonlinéaires par un simple changement de variable. Un exemple est l'équation de BERNOULLI :

$$y' + p(x)y = g(x)y^a,$$

où a est une constante (réelle). On remarque que l'équation est linéaire si $a = 0$ ou $a = 1$.

On utilise la substitution

$$u(x) = (y(x))^{1-a}.$$

Alors

$$\frac{d}{dx}(u(x)) = u'(x) = \frac{d}{dx}((y(x))^{1-a}) = (1-a)(y(x))^{-a} \frac{dy}{dx},$$

c'est-à-dire

$$u' = (1-a)y^{-a}y'.$$

Mais, de l'équation différentielle, on a

$$y' = gy^a - py$$

ce qui implique :

$$\begin{aligned} u' &= (1-a)y^{-a}(gy^a - py) \\ &= (1-a)(g - py^{1-a}) \\ &= (1-a)(g - pu) \\ &= (1-a)g - (1-a)pu. \end{aligned}$$

Nous obtenons donc une équation linéaire en u ,

$$u' + (1-a)p(x)u = (1-a)g(x),$$

qu'on peut résoudre pour u , et revenir à y .

Considérons l'équation de BERNOLLI nonlinéaire

$$y' + y = -\frac{x}{y} \quad \text{ou} \quad y' + (1)y = (-x)y^{-1},$$

avec $p(x) = 1$, $g(x) = -x$ et $a = -1$. La substitution $u(x) = (y(x))^2$ nous donne l'équation linéaire

$$u' + 2u = -2x.$$

Le facteur d'intégration est

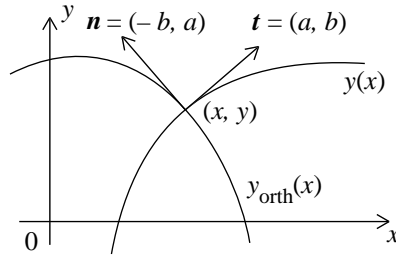
$$\mu(x) = e^{\int 2 dx} = e^{2x}.$$

Alors,

$$\begin{aligned} u(x) &= e^{-2x} \left[\int e^{2x}(-2x) dx + c \right] \\ &= e^{-2x} \left[-x e^{2x} + \frac{1}{2} e^{2x} + c \right] \quad (\text{par intégration par parties}) \\ &= -x + \frac{1}{2} + c e^{-2x}. \end{aligned}$$

Donc la solution générale de l'équation originale est

$$y^2 = \frac{1}{2} - x + c e^{-2x}.$$

FIGURE 1.6. Deux courbes orthogonales au point (x, y) .

1.7. Familles de courbes orthogonales

Une famille de courbes peut être donnée au moyen d'une équation

$$u(x, y) = c,$$

où le paramètre c est explicite, ou d'une équation

$$F(x, y, c) = 0$$

implicite en c .

Dans le 1er cas, les courbes de la famille donnée satisfont l'équation différentielle

$$u_x dx + u_y dy = 0, \quad \text{ou bien} \quad \frac{dy}{dx} = -\frac{u_x}{u_y} = m,$$

où m est la pente de la courbe au point (x, y) . Cette équation différentielle ne contient pas le paramètre c .

Dans le 2ème cas, on a

$$F_x(x, y, c) dx + F_y(x, y, c) dy = 0.$$

Pour éliminer le paramètre c de cette équation différentielle, on résout l'équation $F(x, y, c) = 0$ pour c en fonction de x et y ,

$$c = H(x, y),$$

et l'on substitue cette fonction dans l'équation différentielle

$$\frac{dy}{dx} = -\frac{F_x(x, y, c)}{F_y(x, y, c)} = -\frac{F_x(x, y, H(x, y))}{F_y(x, y, H(x, y))} = m.$$

Soit $\mathbf{t} = (a, b)$ la tangente et $\mathbf{n} = (-b, a)$ la normale à la courbe donnée $y = y(x)$ au point (x, y) de la courbe. Alors la pente de la tangente est

$$y'(x) = \frac{b}{a} = m. \quad (1.23)$$

La pente de la courbe $y_{\text{orth}}(x)$ orthogonale à la courbe $y(x)$ en (x, y) est

$$y'_{\text{orth}}(x) = -\frac{a}{b} = -\frac{1}{m}. \quad (1.24)$$

(V. figure 1.6). Alors, la famille orthogonale satisfait l'équation différentielle

$$y'_{\text{orth}}(x) = -\frac{1}{m(x)}.$$

EXEMPLE 1.16. Soit la famille de cercles

$$x^2 + (y - c)^2 = c^2 \quad (1.25)$$

de centre $(0, c)$ sur l'axe Oy et de rayon $|c|$. Trouver l'équation différentielle de cette famille et celle de la famille orthogonale, résoudre cette dernière équation et tracer quelques courbes des deux familles.

RÉSOLUTION. On obtient l'équation différentielle de la famille donnée en dérivant (1.25) par rapport à x ,

$$2x + 2(y - c)y' = 0 \implies y' = -\frac{x}{y - c}.$$

On résout (1.25) pour c ,

$$x^2 + y^2 - 2yc + c^2 = c^2 \implies c = \frac{x^2 + y^2}{2y},$$

et l'on substitue cette valeur dans l'équation différentielle

$$y' = -\frac{x}{y - \frac{x^2 + y^2}{2y}} = -\frac{2xy}{2y^2 - x^2 - y^2} = \frac{2xy}{x^2 - y^2}.$$

L'équation différentielle de la famille orthogonale est alors

$$y'_{\text{orth}} = -\frac{x^2 - y_{\text{orth}}^2}{2xy_{\text{orth}}}.$$

On réécrit cette équation sous forme différentielle $M dx + N dy = 0$, sans la mention "orth" :

$$(x^2 - y^2) dx + 2xy dy = 0.$$

Puisque $M_y = -2y$ et $N_x = 2y$, cette équation n'est pas exacte, mais

$$\frac{M_y - N_x}{N} = \frac{-2y - 2y}{2xy} = -\frac{2}{x} = f(x)$$

est une fonction de x seulement. Donc

$$\mu(x) = e^{-\int (2/x) dx} = x^{-2}$$

est un facteur d'intégration. On multiplie l'équation différentielle par $\mu(x)$:

$$\left(1 - \frac{y^2}{x^2}\right) dx + 2\frac{y}{x} dy = 0,$$

et l'on résout par la méthode pratique :

$$u(x, y) = \int 2\frac{y}{x} dy + T(x) = \frac{y^2}{x} + T(x),$$

$$u_x(x, y) = -\frac{y^2}{x^2} + T'(x) = 1 - \frac{y^2}{x^2},$$

$$T'(x) = 1 \implies T(x) = x,$$

$$u(x, y) = \frac{y^2}{x} + x = c_1.$$

On a donc la solution générale :

$$x^2 + y^2 = c_1 x;$$

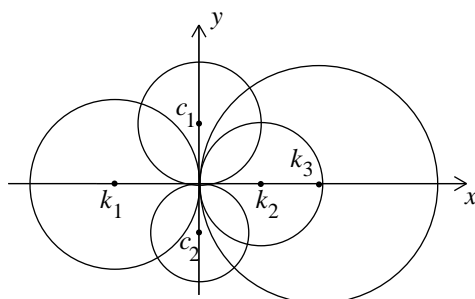


FIGURE 1.7. Quelques courbes des deux familles orthogonales.

c'est une famille de cercles. On réécrit cette solution d'une façon plus explicite :

$$\begin{aligned} x^2 - 2 \frac{c_1}{2} x + \frac{c_1^2}{4} + y^2 &= \frac{c_1^2}{4}, \\ \left(x - \frac{c_1}{2}\right)^2 + y^2 &= \left(\frac{c_1}{2}\right)^2, \\ (x - k)^2 + y^2 &= k^2. \end{aligned}$$

La famille orthogonale est donc une famille de cercles de centre $(k, 0)$ sur l'axe Ox et de rayon $|k|$. La figure 1.7 montre quelques courbes des deux familles. \square

1.8. Champ des tangentes et solutions approchées

On peut utiliser les solutions approchées d'une équation différentielle si l'on ne peut trouver la solution exacte ou si la complexité de la formule de cette solution en rend l'utilisation très difficile. Dans ce cas, on a recours soit à une méthode de résolution numérique (V. chapitre 10), ou soit à la méthode du champ des tangentes. La méthode du champ des tangentes nous permet de tracer plusieurs courbes intégrales, sans pour autant résoudre l'équation différentielle.

La méthode du champ des tangentes s'applique aux équations différentielles de la forme générale

$$y' = f(x, y). \quad (1.26)$$

Il suffit de tracer une courbe intégrale de pente y' . La pente de la courbe qui passe par le point (x_0, y_0) est égale à $f(x_0, y_0)$ en ce point. Ainsi, on peut tracer des petits segments de droite de pente $f(x, y)$ en plusieurs points (x, y) et tracer une courbe intégrale suivant le champ des tangentes.

En pratique, on trace d'abord, des courbes de pentes constantes, $f(x, y) = \text{const}$, appelées *lignes isoclines*, puis on trace le long de chaque ligne isocline $f(x, y) = k$ plusieurs segments de droite de pente k . On obtient ainsi un champ de tangentes. Enfin, on trace une solution approchée de l'équation (1.26).

EXEMPLE 1.17. Tracer le champ des tangentes de l'équation différentielle du 1er ordre

$$y' = xy \quad (1.27)$$

et la solution approchée passant par le point $(1, 2)$.

RÉSOLUTION. Les courbes de pentes constantes sont les hyperboles équilatérales $xy = k$ ainsi que les axes Ox et Oy tracés sur la fig. 1.8 \square

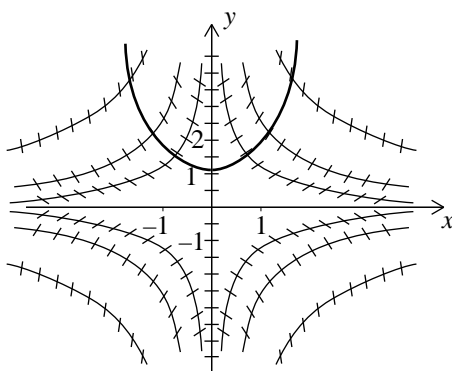


FIGURE 1.8. Champs des tangentes pour l'exemple 1.17.

1.9. Existence et unicité de la solution

DÉFINITION 1.3. Une fonction $f(y)$ est lipschitzienne sur $]c, d[$ s'il existe une constante $M > 0$, appelée constante de Lipschitz, telle que

$$|f(z) - f(y)| \leq M|z - y|, \quad \text{pour tout } y, z \in]c, d[. \quad (1.28)$$

On remarque que la condition (1.28) implique l'existence des dérivées premières à gauche et à droite de $f(y)$, mais non leur égalité. Géométriquement, la pente de la courbe $f(y)$ reste bornée sur $]c, d[$.

On énonce le théorème d'existence et d'unicité suivant.

THÉORÈME 1.3 (Existence et unicité). *Soit le problème à valeur initiale*

$$y' = f(x, y), \quad y(x_0) = y_0. \quad (1.29)$$

Si la fonction $f(x, y)$ est continue et bornée,

$$|f(x, y)| \leq K,$$

sur le rectangle

$$R : |x - x_0| < a, \quad |y - y_0| < b,$$

et lipschitzienne en y sur R , alors (1.29) admet une et une seule solution pour tout x satisfaisant

$$|x - x_0| < \alpha, \quad \text{où } \alpha = \min\{a, b/K\}.$$

On applique le théorème 1.3 à l'exemple suivant.

EXEMPLE 1.18. Résoudre le problème à valeur initiale :

$$yy' + x = 0, \quad y(0) = -2,$$

et tracer la solution.

RÉSOLUTION. **(a) Résolution analytique.**— On écrit l'équation différentielle sous la forme $y' = f(x, y)$,

$$y' = -\frac{x}{y} := f(x, y).$$

Puisque $f(x, y)$ n'est pas continue en $y = 0$, on aura une solution pour $y < 0$ et une autre pour $y > 0$. On sépare l'équation et l'on intègre :

$$\begin{aligned} \int x dx + \int y dy &= 0, \\ \frac{x^2}{2} + \frac{y^2}{2} &= c_1, \\ x^2 + y^2 &= r^2. \end{aligned}$$

La solution générale est donc une famille de cercles de centre l'origine et de rayon $|r|$. On a les deux solutions

$$y_{\pm}(x) = \begin{cases} \sqrt{r^2 - x^2}, & y > 0, \\ -\sqrt{r^2 - x^2}, & y < 0. \end{cases}$$

Puisque $y(0) = -2$, on prend la 2ème solution et l'on détermine la valeur de r :

$$0^2 + (-2)^2 = r^2 \implies r = 2.$$

Donc la solution, qui est unique, est

$$y(x) = -\sqrt{4 - x^2}, \quad -2 < x < 2.$$

On voit que la pente $y'(x)$ de la solution tend vers $\pm\infty$ lorsque $y \rightarrow 0_{\pm}$. Pour avoir une solution continue au voisinage de $y = 0$, on résout pour $x = x(y)$.

(b) Résolution par Matlab symbolique.—

```
dsolve('y*Dy=-x', 'y(0)=-2', 'x')
y = -(-x^2+4)^(1/2)
```

(c) Résolution par Matlab numérique.— La résolution numérique de ce problème à valeur initiale est un peu futée parce que la solution générale admet deux branches : y_{\pm} . On a besoin d'une fonction fichier M pour implémenter le solveur `ode23`. Le fichier M `halfcircle.m` est

```
function yprime = halfcircle(x,y);
yprime = -x/y;
```

Pour obtenir la branche inférieure de la solution générale, on appelle le solveur `ode23` et la commande `plot` :

```
xspan1 = [0 -2]; % span de x = 0 = 0 a x = -2
xspan2 = [0 2]; % span de x = 0 a x = 2
y0 = [0; -2]; % condition initiale
[x1,y1] = ode23('halfcircle',xspan1,y0);
[x2,y2] = ode23('halfcircle',xspan2,y0);
plot(x1,y1(:,2),x2,y2(:,2))
axis('equal')
xlabel('x')
ylabel('y')
title('Trace de la solution')
```

La solution numérique est tracée dans la figure 1.9. □

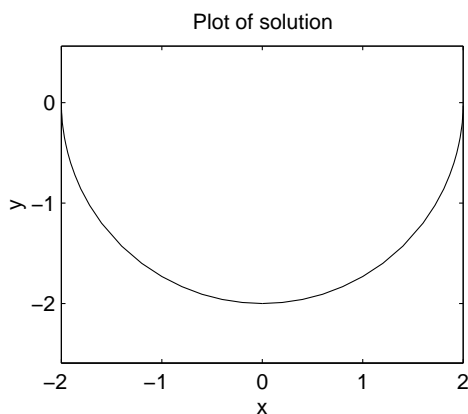


FIGURE 1.9. Graphe de la solution pour l'exemple 1.18.

Sous les hypothèses du théorème 1.3, la solution du problème (1.29) peut s'obtenir au moyen de la méthode de Picard, c'est-à-dire la suite $y_0, y_1, \dots, y_n, \dots$, définie par la récurrence de Picard,

$$y_n(x) = y_0 + \int_{x_0}^x f(t, y_{n-1}(t)) dt, \quad n = 1, 2, \dots, \quad (1.30)$$

converge vers la solution $y(x)$.

On emploie la récurrence de Picard et la méthode de la section 1.6 pour résoudre la même équation dans les deux exemples suivants. À l'exemple 6.4, on va résoudre la même équation au moyen d'une série de puissances. Les trois méthodes produisent la même solution série. De plus, à l'exemple 10.9, on résoudra cette équation par une méthode numérique.

EXEMPLE 1.19. Résoudre le problème à valeur initiale :

$$y' = xy + 1, \quad y(0) = 1,$$

par la récurrence de Picard.

RÉSOLUTION. Puisque la dérivée partielle en y du 1er ordre de la fonction $f(x, y) = 1 + xy$:

$$\partial_y f(x, y) = x,$$

est bornée sur un intervalle borné $x_0 \leq x \leq a < \infty$, la récurrence de Picard (1.30),

$$y_n(x) = y_0 + \int_{x_0}^x f(t, y_{n-1}(t)) dt, \quad n = 1, 2, \dots,$$

converge vers la solution $y(x)$. On prend $x_0 = 0$ et $y_0 = 1$. Alors,

$$\begin{aligned} y_1(x) &= 1 + \int_0^x (1+t) dt \\ &= 1 + x + \frac{x^2}{2}, \\ y_2(x) &= 1 + \int_0^x \left(1 + t + t^2 + \frac{t^3}{2}\right) dt \\ &= 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{8}, \\ y_3(x) &= 1 + \int_0^x (1 + ty_2(t)) dt, \end{aligned}$$

et ainsi de suite. □

EXEMPLE 1.20. Résoudre le problème à valeur initiale :

$$y' - xy = 1, \quad y(0) = 1,$$

par la méthode de la section 1.6 pour les équations linéaires du 1er ordre.

RÉSOLUTION. Le facteur d'intégration :

$$\mu(x) = e^{-\int x dx} = e^{-x^2/2}.$$

transforme le 1er membre en une dérivée totale. On multiplie l'équation par $\mu(x)$:

$$\frac{d}{dx} \left(e^{-x^2/2} y \right) = e^{-x^2/2},$$

et l'on intègre de 0 à x :

$$e^{-x^2/2} y(x) = \int_0^x e^{-t^2/2} dt + c.$$

On prend $x = 0$ et $y(0) = 1$. D'où il suit que $c = 1$. Alors,

$$y(x) = e^{x^2/2} \left[1 + \int_0^x e^{-t^2/2} dt \right].$$

Puisque l'on ne peut faire l'intégration sous forme algébrique, on développe les deux exponentielles en séries de puissances convergentes. On intègre la seconde série terme à terme et l'on multiplie les deux séries terme à terme :

$$\begin{aligned} y(x) &= e^{x^2/2} \left[1 + \int_0^x \left(1 - \frac{t^2}{2} + \frac{t^4}{8} - \frac{t^6}{48} + \dots \right) dt \right] \\ &= e^{x^2/2} \left(1 + x - \frac{x^3}{6} + \frac{x^5}{40} - \frac{x^7}{336} + \dots \right) \\ &= \left(1 + \frac{x^2}{2} + \frac{x^4}{8} + \frac{x^6}{48} + \dots \right) \left(1 + x - \frac{x^3}{6} + \frac{x^5}{40} - \frac{x^7}{336} + \dots \right) \\ &= 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{8} + \dots \end{aligned}$$

On obtient la même solution série qu'à l'exemple 1.19.

La commande `dsolve` de Matlab symbolique nous donne la solution en terme de la fonction de l'erreur `erf(x)` de Maple :

```
>> dsolve('Dy=x*y+1', 'y(0)=1', 'x')
y=1/2*exp(1/2*x^2)*pi^(1/2)*2^(1/2)*erf(1/2*2^(1/2)*x)+exp(1/2*x^2)
```

comme on pouvait s'y attendre. \square

L'exemple suivant montre que la simple continuité de la fonction $f(x, y)$ n'assure pas l'unicité de la solution de l'équation $y' = f(x, y)$ si f n'est pas lipschitzienne par rapport à y .

EXEMPLE 1.21. Montrer que le problème à valeur initiale :

$$y' = 3y^{2/3}, \quad y(x_0) = y_0,$$

admet plusieurs solutions.

RÉSOLUTION. Le second membre de l'équation est continu pour tout y et, puisqu'il est indépendant de x , il est continu sur tout le plan xy . Cependant il n'est pas lipschitzien par rapport à y en $y = 0$ puisque $f_y(x, y) = 2y^{-1/3}$ n'est même pas définie en $y = 0$. On voit que $y(x) \equiv 0$ est solution de l'équation différentielle. De plus, pour $a \leq b$,

$$y(x) = \begin{cases} (x-a)^3, & t < a, \\ 0, & a \leq x \leq b, \\ (x-b)^3, & t > b, \end{cases}$$

est aussi une solution. On peut satisfaire la condition initiale par un choix approprié d'un des paramètres a ou b . On obtient une famille de solution en faisant varier l'autre paramètre. Donc la solution n'est pas unique. \square

Équations différentielles linéaires d'ordre deux

Dans ce chapitre, on introduit quelques concepts fondamentaux sur les équations différentielles linéaires du second ordre. On résout les équations à coefficients constants et l'équation d'Euler–Cauchy.

Au chapitre suivant, on reprendra ces idées pour les équations linéaires non homogènes d'ordre quelconque.

2.1. Équations linéaires homogènes

Soit l'équation différentielle linéaire non homogène d'ordre deux :

$$y'' + f(x)y' + g(x)y = r(x). \quad (2.1)$$

L'équation est linéaire en y , y' et y'' . Elle est non homogène si le 2ème membre, $r(x)$, est non nul.

On représentera souvent un opérateur différentiel linéaire quelconque par la lettre L :

$$L := a_n(x)D^n + a_{n-1}(x)D^{n-1} + \dots + a_1(x)D + a_0(x), \quad D = ' = \frac{d}{dx}.$$

Si le 2ème membre de (2.1) est nul, on a une équation *homogène* :

$$Ly := y'' + f(x)y' + g(x)y = 0. \quad (2.2)$$

THÉORÈME 2.1. *Les solutions de l'équation homogène (2.2) forment un espace vectoriel.*

DÉMONSTRATION. Soit y_1 et y_2 deux solutions de (2.2). La linéarité de L implique:

$$L(\alpha y_1 + \beta y_2) = \alpha Ly_1 + \beta Ly_2 = 0, \quad \alpha, \beta \in \mathbb{R}. \quad \square$$

2.2. Équations homogènes à coefficients constants

Soit l'équation différentielle linéaire homogène du second ordre à *coefficients constants* :

$$y'' + ay' + by = 0. \quad (2.3)$$

On résout cette équation en supposant que la solution est de la forme exponentielle suivante :

$$y = e^{\lambda x}.$$

Alors,

$$\lambda^2 e^{\lambda x} + a\lambda e^{\lambda x} + b e^{\lambda x} = 0, \quad (2.4)$$

$$e^{\lambda x} (\lambda^2 + a\lambda + b) = 0. \quad (2.5)$$

Puisque $e^{\lambda x}$ ne s'annule jamais, on obtient l'équation caractéristique

$$\lambda^2 + a\lambda + b = 0 \quad (2.6)$$

pour λ et les *valeurs propres*

$$\lambda_{1,2} = \frac{-a \pm \sqrt{a^2 - 4b}}{2}. \quad (2.7)$$

Si $\lambda_1 \neq \lambda_2$, on a deux solutions distinctes

$$y_1 = e^{\lambda_1 x}, \quad y_2 = e^{\lambda_2 x},$$

et la *solution générale*, qui contient deux constantes arbitraires, est

$$y = c_1 y_1 + c_2 y_2.$$

EXEMPLE 2.1. Résoudre l'équation différentielle

$$y'' + 5y' + 6y = 0.$$

RÉSOLUTION. L'équation caractéristique est

$$\lambda^2 + 5\lambda + 6 = (\lambda + 2)(\lambda + 3) = 0.$$

Alors $\lambda_1 = -2$ et $\lambda_2 = -3$. La solution générale est donc :

$$y = c_1 e^{-2x} + c_2 e^{-3x}. \quad \square$$

2.3. Base de l'espace solution

On étend aux fonctions définies sur $[a, b]$ la notion d'indépendance linéaire de deux vecteurs dans \mathbb{R}^2 .

DÉFINITION 2.1. Deux fonctions $f_1(x)$ et $f_2(x)$ sont linéairement indépendantes sur $[a, b]$ si l'identité

$$c_1 f_1(x) + c_2 f_2(x) \equiv 0 \quad \text{sur } [a, b] \quad (2.8)$$

implique

$$c_1 = c_2 = 0.$$

Sinon, elles sont linéairement dépendantes.

Si $f_1(x)$ et $f_2(x)$ sont linéairement dépendantes sur $[a, b]$, il existe deux nombres $(c_1, c_2) \neq (0, 0)$ tels que l'identité (2.8) est satisfaite sur $[a, b]$. Supposant que $c_1 \neq 0$, on a

$$\frac{f_1(x)}{f_2(x)} \equiv -\frac{c_2}{c_1} = \text{const.} \quad (2.9)$$

On conclut que si

$$\frac{f_1(x)}{f_2(x)} \neq \text{const.} \quad \text{sur } [a, b], \quad (2.10)$$

alors f_1 et f_2 sont linéairement indépendantes sur $[a, b]$. On emploiera souvent cette caractérisation d'indépendance linéaire pour deux fonctions.

DÉFINITION 2.2. La solution générale de l'équation homogène (2.2) engendre l'espace vectoriel des solutions de (2.2).

THÉORÈME 2.2. Soit $y_1(x)$ et $y_2(x)$ deux solutions de (2.2) sur $[a, b]$. Alors, la solution

$$y(x) = c_1 y_1(x) + c_2 y_2(x)$$

est une solution générale de (2.2) si et seulement si y_1 et y_2 sont linéairement indépendantes sur $[a, b]$.

DÉMONSTRATION. On donnera la démonstration pour les équations d'ordre n quelconque au chapitre suivant. \square

L'exemple suivant illustre l'utilité de la solution générale.

EXEMPLE 2.2. Résoudre le problème aux valeurs initiales :

$$y'' + y' - 2y = 0, \quad y(0) = 4, \quad y'(0) = 1.$$

RÉSOLUTION. (a) **Résolution analytique.**— L'équation caractéristique est

$$\lambda^2 + \lambda - 2 = (\lambda - 1)(\lambda + 2) = 0.$$

Alors $\lambda_1 = 1$ et $\lambda_2 = -2$. Les solutions

$$y_1(x) = e^x, \quad y_2(x) = e^{-2x}$$

sont linéairement indépendantes car

$$\frac{y_1(x)}{y_2(x)} = e^{3x} \neq \text{const.}$$

La solution générale est donc

$$y = c_1 e^x + c_2 e^{-2x}.$$

On détermine la valeur des constantes au moyen des conditions initiales :

$$\begin{aligned} y(0) &= c_1 + c_2 = 4, \\ y'(x) &= c_1 e^x - 2c_2 e^{-2x}, \\ y'(0) &= c_1 - 2c_2 = 1. \end{aligned}$$

On obtient donc le système linéaire

$$\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad \text{c.-à-d.} \quad A\mathbf{c} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}.$$

Puisque

$$\det A = -3 \neq 0,$$

la solution \mathbf{c} est unique. On obtient cette solution par la règle de Cramer :

$$c_1 = \frac{1}{-3} \begin{vmatrix} 4 & 1 \\ 1 & -2 \end{vmatrix} = \frac{-9}{-3} = 3, \quad c_2 = \frac{1}{-3} \begin{vmatrix} 1 & 4 \\ 1 & 1 \end{vmatrix} = \frac{-3}{-3} = 1.$$

La solution du problème aux valeurs initiales est donc

$$y(x) = 3e^x + e^{-2x}.$$

Cette solution est unique.

(b) **Résolution par Matlab symbolique.**—

```
dsolve('D2y+Dy-2*y=0', 'y(0)=4', 'Dy(0)=1', 'x')
y = 3*exp(x)+exp(-2*x)
```

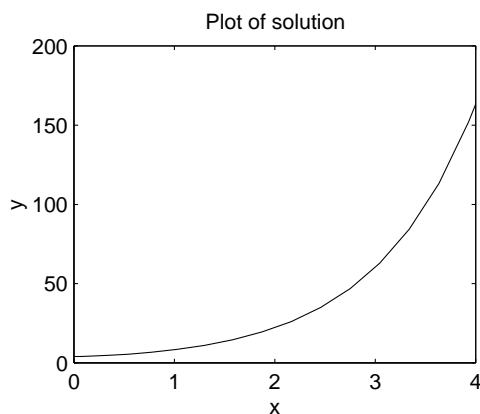


FIGURE 2.1. Graphe de la solution de l'équation linéaire de l'exemple 2.2.

(c) **Résolution par Matlab numérique.**— On récrit l'équation différentielle du second ordre en un système du premier ordre au moyen des variables

$$\begin{aligned}y_1 &= y, \\ y_2 &= y',\end{aligned}$$

Alors,

$$\begin{aligned}y_1' &= y_2, \\ y_2' &= 2y_1 - y_2.\end{aligned}$$

Le fichier M `exp22.m` :

```
function yprime = exp22(x,y);
yprime = [y(2); 2*y(1)-y(2)];
```

On appelle le solveur `ode23` et la commande `plot` :

```
xspan = [0 4]; % solution sur 0<=x<=4
y0 = [4; 1]; % conditions initiales
[x,y] = ode23('exp22',xspan,y0);
subplot(2,2,1); plot(x,y(:,1))
```

La solution numérique se trouve à la fig. 2.1. □

2.4. Solutions indépendantes

La forme des solutions indépendantes de l'équation homogène

$$Ly := y'' + ay' + by = 0 \tag{2.11}$$

dépend de la forme des racines

$$\lambda_{1,2} = \frac{-a \pm \sqrt{a^2 - 4b}}{2} \tag{2.12}$$

de l'équation caractéristique

$$\lambda^2 + a\lambda + b = 0. \tag{2.13}$$

Soit $\Delta = a^2 - 4b$ le discriminant de l'équation (2.13). Il faut considérer trois cas : $\lambda_1 \neq \lambda_2$ réelles si $\Delta > 0$, $\lambda_2 = \bar{\lambda}_1$ complexes si $\Delta < 0$ et $\lambda_1 = \lambda_2$ réelles si $\Delta = 0$.

Cas I. Dans le cas de deux valeurs propres réelles et distinctes, $\lambda_1 \neq \lambda_2$, on a vu à la section 2.3 que les deux solutions

$$y_1 = e^{\lambda_1 x}, \quad y_2 = e^{\lambda_2 x},$$

sont indépendantes. La solution générale est donc

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}. \quad (2.14)$$

Cas II. Dans le cas de deux valeurs propres complexes, conjuguées l'une de l'autre, on a

$$\lambda_1 = \alpha + i\beta, \quad \lambda_2 = \alpha - i\beta = \bar{\lambda}_1, \quad \text{où } i = \sqrt{-1}.$$

On emploie l'identité d'Euler,

$$e^{i\theta} = \cos \theta + i \sin \theta. \quad (2.15)$$

On a alors les deux solutions complexes :

$$\begin{aligned} u_1(x) &= e^{(\alpha+i\beta)x} = e^{\alpha x} (\cos \beta x + i \sin \beta x), \\ u_2(x) &= e^{(\alpha-i\beta)x} = e^{\alpha x} (\cos \beta x - i \sin \beta x) = \overline{u_1(x)}. \end{aligned}$$

Puisque $\lambda_1 \neq \lambda_2$, les solutions u_1 et u_2 sont indépendantes. Pour avoir deux solutions réelles indépendantes, on fait le changement de base suivant, ou, de façon équivalente, on prend les parties réelle et imaginaire de u_1 puisque a et b sont réels et l'équation (2.11) est homogène (en effet, les parties réelle et imaginaire d'une solution complexe d'une équation linéaire homogène à coefficients réels sont aussi des solutions). Alors

$$y_1(x) = \Re u_1(x) = \frac{1}{2}[u_1(x) + u_2(x)] = e^{\alpha x} \cos \beta x, \quad (2.16)$$

$$y_2(x) = \Im u_1(x) = \frac{1}{2i}[u_1(x) - u_2(x)] = e^{\alpha x} \sin \beta x. \quad (2.17)$$

On voit immédiatement que y_1 et y_2 sont indépendantes. Alors, la solution générale est

$$y(x) = c_1 e^{\alpha x} \cos \beta x + c_2 e^{\alpha x} \sin \beta x. \quad (2.18)$$

Cas III. Dans le cas d'une valeur propre réelle double on a

$$\lambda = \lambda_1 = \lambda_2 = -\frac{a}{2}$$

et (2.11) admet une solution de la forme

$$y_1(x) = e^{\lambda x}. \quad (2.19)$$

Pour obtenir une seconde solution indépendante de y_1 par variation des paramètres, on pose

$$y_2(x) = u(x)y_1(x). \quad (2.20)$$

Il est important de remarquer que le paramètre u est fonction de x et que y_1 est solution de (2.11). On substitue y_2 dans (2.11) et l'on additionne les trois équations suivantes :

$$\begin{aligned} by_2(x) &= bu(x)y_1(x) \\ ay_2'(x) &= au(x)y_1'(x) + ay_1(x)u'(x) \\ y_2''(x) &= u(x)y_1''(x) + 2y_1'(x)u'(x) + y_1(x)u''(x) \\ Ly_2 &= u(x)Ly_1 + [ay_1(x) + 2y_1'(x)]u'(x) + y_1(x)u''(x). \end{aligned}$$

Le 1er membre de la somme est nul puisqu'on suppose que y_2 est solution de $Ly = 0$. Le premier terme du 2ème membre est nul puisque y_1 est solution de $Ly = 0$.

Le 2ème terme du 2ème membre est nul parce que

$$\lambda = -\frac{a}{2} \in \mathbb{R}$$

et $y_1'(x) = \lambda y_1(x)$, c'est-à-dire

$$ay_1(x) + 2y_1'(x) = a e^{-ax/2} - a e^{-ax/2} = 0.$$

Il suit que

$$u''(x) = 0,$$

d'où

$$u'(x) = k_1$$

et

$$u(x) = k_1x + k_2.$$

On a donc

$$y_2(x) = k_1x e^{\lambda x} + k_2 e^{\lambda x}.$$

Il suffit de prendre $k_2 = 0$ parce que le 2ème terme du 2ème membre est déjà contenu dans l'enveloppe linéaire de y_1 . On peut aussi prendre $k_1 = 1$ puisque la solution générale contient déjà une constante multipliant y_2 .

On voit immédiatement que les solutions

$$y_1(x) = e^{\lambda x}, \quad y_2(x) = x e^{\lambda x},$$

sont linéairement indépendantes.

Alors, la solution générale est

$$y(x) = c_1 e^{\lambda x} + c_2 x e^{\lambda x}. \quad (2.21)$$

EXEMPLE 2.3. On considère les trois problèmes suivants.

i) Trouver la solution générale de l'équation homogène à coefficients constants :

$$y'' - 7y' + 12y = 0.$$

L'équation caractéristique est

$$\lambda^2 - 7\lambda + 12 = (\lambda - 3)(\lambda - 4) = 0.$$

Donc $\lambda_1 = 3$ et $\lambda_2 = 4$ (Cas I). Alors, la solution générale est

$$y(x) = c_1 e^{3x} + c_2 e^{4x}.$$

ii) Trouver la solution générale de l'équation homogène à coefficients constants :

$$y'' + 4y' + 10y = 0.$$

L'équation caractéristique est

$$\lambda^2 + 4\lambda + 10 = 0.$$

Donc

$$\lambda_{1,2} = \frac{-4 \pm \sqrt{(4)^2 - 4(10)}}{2} = \frac{-4 \pm \sqrt{-24}}{2} = -2 \pm i\sqrt{6}.$$

C'est le Cas II et la solution générale est

$$y(x) = c_1 e^{-2x} \cos(\sqrt{6}x) + c_2 e^{-2x} \sin(\sqrt{6}x).$$

iii) Résoudre le problème à valeur initiale :

$$y'' - 4y' + 4y = 0, \quad y(0) = 0, \quad y'(0) = 3.$$

L'équation caractéristique est

$$\lambda^2 - 4\lambda + 4 = (\lambda - 2)^2 = 0.$$

Donc $\lambda_1 = \lambda_2 = 2$ (Cas III) et la solution générale est

$$y(x) = c_1 e^{2x} + c_2 x e^{2x}.$$

Par la première condition initiale,

$$y(0) = c_1 e^0 + c_2 0 e^0 = c_1 = 0.$$

Donc $y(x) = c_2 x e^{2x}$. Alors

$$y'(x) = c_2 e^{2x} + 2c_2 x e^{2x}$$

et, en $x = 0$,

$$y'(0) = c_2 e^0 + 2c_2 0 e^0 = c_2 = 3.$$

La solution unique est

$$y(x) = 3x e^{2x}.$$

2.5. Modélisation en mécanique

On considère quelques modèles de la mécanique élémentaire.

EXEMPLE 2.4 (Oscillation libre). Soit une ressort en position verticale pendante fixé à une poutre rigide. Le ressort résiste à l'extension et à la compression et sa constante de restauration (constante de Hooke) est k . Etudier le problème de l'oscillation libre verticale d'une masse de m kg fixée au bout inférieur du ressort.

RÉSOLUTION. On considère la direction Oy vers le bas comme positive. Soit s_0 m l'étirement du ressort par le poids au repos en position $y = 0$ (V. figure 2.2).

On néglige toute friction. La force due à la gravité est

$$F_1 = mg, \quad \text{où } g = 9.8 \text{ m/sec}^2.$$

La force de restauration exercée par le ressort est

$$F_2 = -k s_0.$$

Quand le système est au repos, la résultante est nulle,

$$F_1 + F_2 = 0,$$

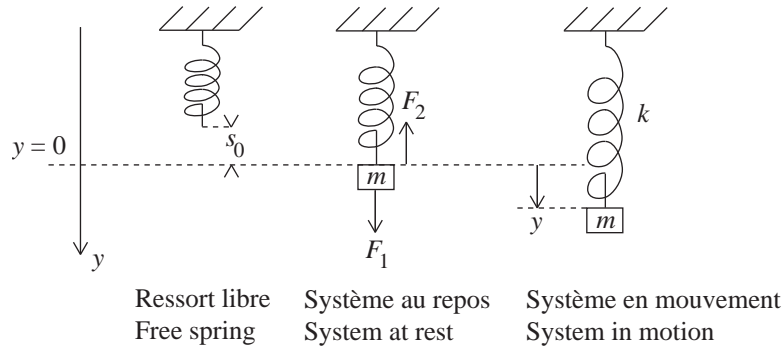


FIGURE 2.2. Système non amorti.

par la 2ème loi de Newton. Puisque le système en mouvement est non amorti, par la même loi, la résultante est

$$m a = -k y.$$

Or l'accélération est donnée par $a = y''$. Donc

$$m y'' + k y = 0, \quad \text{ou bien} \quad y'' + \omega^2 y = 0, \quad \omega = \sqrt{\frac{k}{m}},$$

où $\omega/2\pi$ Hz est la fréquence du système. L'équation caractéristique de cette équation différentielle,

$$\lambda^2 + \omega^2 = 0,$$

admet les valeurs propres imaginaires

$$\lambda_{1,2} = \pm i\omega.$$

La solution générale est donc

$$y(t) = c_1 \cos \omega t + c_2 \sin \omega t.$$

On voit que le système oscille librement sans perte d'énergie. □

L'**amplitude**, A , et la **période**, p , de ce système sont

$$A = \sqrt{c_1^2 + c_2^2}, \quad p = \frac{2\pi}{\omega}.$$

En effet, si φ est le déphasage, on a

$$\begin{aligned} y(t) &= A \cos(\omega t + \varphi) \\ &= A \cos \varphi \cos \omega t - A \sin \varphi \sin \omega t \\ &= c_1 \cos \omega t + c_2 \sin \omega t. \end{aligned}$$

Alors,

$$c_1^2 + c_2^2 = (A \cos \varphi)^2 + (A \sin \varphi)^2 = A^2.$$

EXEMPLE 2.5 (Système amorti). Soit une ressort en position verticale pendante fixé à une poutre rigide. Le ressort résiste à l'extension et à la compression et sa constante de restauration (constante de Hooke) est k . Étudier le problème du mouvement vertical amorti d'une masse de m kg fixée au bout inférieur du ressort (V. figure 2.3). La constante d'amortissement est c .

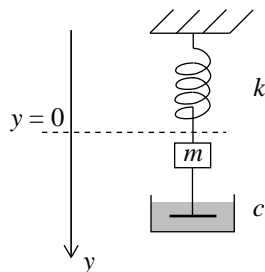


FIGURE 2.3. Système amorti.

RÉSOLUTION. On considère la direction Oy vers le bas comme positive. Soit s_0 m l'étirement du ressort par le poids au repos en position $y = 0$ (V. figure 2.2).

La force due à la gravité est

$$F_1 = mg, \quad \text{où } g = 9.8 \text{ m/sec}^2.$$

La force de restauration exercée par le ressort est

$$F_2 = -k s_0.$$

Quand le système est au repos, la résultante est nulle,

$$F_1 + F_2 = 0,$$

par la 2ème loi de Newton. Puisque l'amortissement est dans la direction opposée au mouvement, par la même loi, la résultante est

$$m a = -c y' - k y.$$

Or l'accélération est donnée par $a = y''$. Donc

$$m y'' + c y' + k y = 0, \quad \text{ou bien } y'' + \frac{c}{m} y' + \frac{k}{m} y = 0.$$

L'équation caractéristique de cette équation différentielle,

$$\lambda^2 + \frac{c}{m} \lambda + \frac{k}{m} = 0,$$

admet les valeurs propres

$$\lambda_{1,2} = -\frac{c}{2m} \pm \frac{1}{2m} \sqrt{c^2 - 4mk} =: -\alpha \pm \beta, \quad \alpha > 0.$$

On a les trois cas suivants.

Cas I : Sur-amortissement. Si $c^2 > 4mk$, le système est sur-amorti. Les deux valeurs propres sont négatives puisque

$$\lambda_1 = -\frac{c}{2m} - \frac{1}{2m} \sqrt{c^2 - 4mk} < 0, \quad \lambda_1 \lambda_2 = \frac{k}{m} > 0.$$

La solution générale,

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t},$$

décroit exponentiellement vers zéro sans oscillation du système.

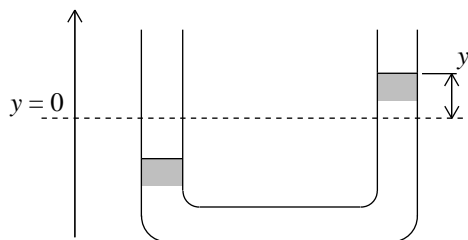


FIGURE 2.4. Mouvement vertical d'un liquide dans un tube en U.

Cas II : Sous-amortissement. Si $c^2 < 4mk$, le système est sous-amorti. Les deux valeurs propres sont complexes, conjuguées l'une de l'autre,

$$\lambda_{1,2} = -\frac{c}{2m} \pm \frac{i}{2m} \sqrt{4mk - c^2} =: -\alpha \pm i\beta, \quad \alpha > 0.$$

La solution générale,

$$y(t) = c_1 e^{-\alpha t} \cos \beta t + c_2 e^{-\alpha t} \sin \beta t,$$

décroît exponentiellement vers zéro avec oscillation du système.

Cas III : Amortissement critique. Si $c^2 = 4mk$, le système est critique-ment amorti. Les deux valeurs propres sont réelles et égales,

$$\lambda_{1,2} = -\frac{c}{2m} = -\alpha, \quad \alpha > 0.$$

La solution générale,

$$y(t) = c_1 e^{-\alpha t} + c_2 t e^{-\alpha t} = (c_1 + c_2 t) e^{-\alpha t},$$

décroît exponentiellement vers zéro avec une augmentation initiale de $y(t)$ si $c_2 > 0$. \square

EXEMPLE 2.6 (Oscillation aqueuse dans un tube en forme de U).

Calculer la fréquence du mouvement oscillatoire vertical de 2 L d'eau dans un tube en forme de U de 0.04 m de diamètre.

RÉSOLUTION. On néglige la friction entre le liquide et la paroi du tube. La masse du liquide est $m = 2$ kg. Le volume causant la force de restauration est

$$\begin{aligned} V &= \pi r^2 h = \pi(0.02)^2 2y \text{ m}^3 \\ &= \pi(0.02)^2 2000y \text{ L} \end{aligned}$$

(V. figure 2.4). La masse du volume V est

$$M = \pi(0.02)^2 2000y \text{ kg}$$

et la force de restauration est

$$Mg = \pi(0.02)^2 9.8 \times 2000y \text{ N}, \quad g = 9.8 \text{ m/s}^2.$$

Par la 2ème loi de Newton,

$$m y'' = -Mg,$$

c'est-à-dire

$$y'' + \frac{\pi(0.02)^2 9.8 \times 2000}{2} y = 0, \quad \text{ou bien} \quad y'' + \omega_0^2 y = 0,$$

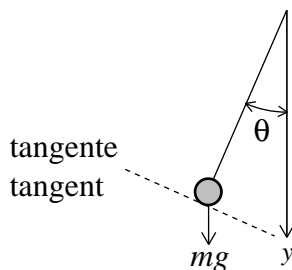


FIGURE 2.5. Pendule en mouvement.

où

$$\omega_0^2 = \frac{\pi(0.02)^2 9.8 \times 2000}{2} = 12.3150.$$

La fréquence est donc

$$\frac{\omega_0}{2\pi} = \frac{\sqrt{12.3150}}{2\pi} = 0.5585 \text{ Hz. } \square$$

EXEMPLE 2.7 (Oscillation d'un pendule). Calculer la fréquence des oscillations de faible amplitude d'un pendule de masse m kg et de longueur $L = 1$ m.

RÉSOLUTION. On néglige la résistance de l'air et la masse de la tige. Soit θ l'angle, en radians, porté par le pendule à partir de la verticale (V. figure 2.5).

La force tangentielle est

$$m a = mL\theta''.$$

Puisque la longueur de la tige est fixée, la composante de la force orthogonale au mouvement est nulle. Il suffit donc de considérer la composante tangentielle de la force de restauration due à la gravité. On a donc

$$mL\theta'' = -mg \sin \theta \approx -mg\theta, \quad g = 9.8,$$

puisque $\sin \theta \approx \theta$ si θ est suffisamment petit. Alors,

$$\theta'' + \frac{g}{L} \theta = 0, \quad \text{ou bien} \quad \theta'' + \omega_0^2 \theta = 0, \quad \text{où} \quad \omega_0^2 = \frac{g}{L} = 9.8.$$

La fréquence est donc

$$\frac{\omega_0}{2\pi} = \frac{\sqrt{9.8}}{2\pi} = 0.498 \text{ Hz. } \square$$

□

2.6. Équations d'Euler-Cauchy

Considérons l'équation d'Euler-Cauchy homogène

$$Ly := x^2 y'' + axy' + by = 0, \quad x > 0. \quad (2.22)$$

A cause de la forme particulière de l'opérateur différentiel à coefficients variables,

$$L = x^2 D^2 + axD + bI, \quad D = ' = \frac{d}{dx},$$

dont chaque terme est de la forme $a_k x^k D^k$, où a_k est une constante, on peut résoudre (2.22) en posant

$$y = x^m \quad (2.23)$$

dans (2.22) :

$$m(m-1)x^m + amx^m + bx^m = x^m[m(m-1) + am + b] = 0.$$

On peut diviser par x^m . On a alors l'équation caractéristique

$$m^2 + (a-1)m + b = 0. \quad (2.24)$$

Les valeurs propres sont

$$m_{1,2} = \frac{1-a}{2} \pm \frac{1}{2}\sqrt{(a-1)^2 - 4b}. \quad (2.25)$$

Il faut considérer trois cas : $m_1 \neq m_2$ réelles, m_1 et $m_2 = \bar{m}_1$ complexes distinctes et $m_1 = m_2$ réelles.

Cas I. Si l'on a deux racines réelles distinctes, la solution générale de (2.22) est

$$y(x) = c_1x^{m_1} + c_2x^{m_2}. \quad (2.26)$$

Cas II. Si l'on a deux racines complexes, conjuguées l'une de l'autre,

$$m_1 = \alpha + i\beta, \quad m_2 = \alpha - i\beta, \quad \beta \neq 0,$$

on a deux solutions complexes indépendantes :

$$u_1 = x^\alpha x^{i\beta} = x^\alpha e^{i\beta \ln x} = x^\alpha [\cos(\beta \ln x) + i \sin(\beta \ln x)]$$

et

$$u_2 = x^\alpha x^{-i\beta} = x^\alpha e^{-i\beta \ln x} = x^\alpha [\cos(\beta \ln x) - i \sin(\beta \ln x)].$$

Pour $x > 0$, on obtient les deux solutions réelles indépendantes suivantes en prenant la partie réelle et la partie imaginaire de u_1 (ou de u_2),

$$y_1(x) = x^\alpha \cos(\beta \ln x), \quad y_2(x) = x^\alpha \sin(\beta \ln x).$$

La solution générale de (2.22) est donc

$$y(x) = c_1x^\alpha \cos(\beta \ln x) + c_2x^\alpha \sin(\beta \ln x). \quad (2.27)$$

Cas III. Si l'on a deux racines réelles égales,

$$m = m_1 = m_2 = \frac{1-a}{2},$$

on a une solution de la forme

$$y_1(x) = x^m.$$

Pour obtenir une 2ème solution indépendante au moyen de la variation des paramètres, on pose

$$y_2 = u(x)y_1(x)$$

dans (2.22) et l'on additionne les 1ers et 2èmes membres respectifs des trois expressions suivantes :

$$\begin{aligned} by_2(x) &= bu(x)y_1(x) \\ axy_2'(x) &= axu(x)y_1'(x) + axy_1(x)u'(x) \\ x^2y_2''(x) &= x^2u(x)y_1''(x) + 2x^2y_1'(x)u'(x) + x^2y_1(x)u''(x) \\ Ly_2 &= u(x)Ly_1 + [axy_1(x) + 2x^2y_1'(x)]u'(x) + x^2y_1(x)u''(x). \end{aligned}$$

Le 1er membre de la somme est nul puisqu'on suppose que y_2 est solution de $Ly = 0$. Le premier terme du 2ème membre est nul puisque y_1 est solution de $Ly = 0$.

Le coefficient de u' est

$$\begin{aligned} axy_1(x) + 2x^2y_1'(x) &= axx^m + 2mx^2x^{m-1} = ax^{m+1} + 2mx^{m+1} \\ &= (a + 2m)x^{m+1} = \left(a + 2\frac{1-a}{2}\right)x^{m+1} = x^{m+1}. \end{aligned}$$

On obtient donc

$$x^2y_1(x)u'' + x^{m+1}u' = x^{m+1}(xu'' + u') = 0, \quad x > 0.$$

On peut diviser par x^{m+1} :

$$xu'' + u' = 0.$$

Puisque u est absente de cette équation différentielle, on peut en réduire l'ordre en posant

$$v = u', \quad v' = u''.$$

Alors on a l'équation séparable

$$x \frac{dv}{dx} + v = 0, \quad \text{c'est-à-dire} \quad \frac{dv}{v} = -\frac{dx}{x},$$

qu'on peut intégrer :

$$\ln |v| = \ln x^{-1} \implies u' = v = \frac{1}{x} \implies u = \ln x.$$

La 2ème solution, indépendante de la 1ère, est

$$y_2 = (\ln x)x^m.$$

La solution générale de (2.22) est donc

$$y(x) = c_1x^m + c_2(\ln x)x^m. \quad (2.28)$$

EXEMPLE 2.8. Trouver la solution générale de l'équation d'Euler-Cauchy

$$x^2y'' - 6y = 0.$$

RÉSOLUTION. (a) **Résolution analytique.**— En posant $y = x^m$ dans l'équation différentielle, on obtient

$$m(m-1)x^m - 6x^m = 0.$$

L'équation caractéristique est

$$m^2 - m - 6 = (m-3)(m+2) = 0.$$

Les valeurs propres,

$$m_1 = 3, \quad m_2 = -2,$$

sont réelles et distinctes. La solution générale est donc

$$y(x) = c_1x^3 + c_2x^{-2}.$$

(b) **Résolution par Matlab symbolique.**—

```
dsolve('x^2*D2y=6*y', 'x')
y = (C1+C2*x^5)/x^2
```

□

EXEMPLE 2.9. Résoudre :

$$x^2y'' - 6y = 0, \quad y(1) = 2, \quad y'(1) = 1.$$

RÉSOLUTION. (a) **Résolution analytique.**— La solution générale trouvée à l'exemple 2.8 est

$$y(x) = c_1x^3 + c_2x^{-2}.$$

Utilisant les conditions initiales, on obtient le système linéaire en c_1 et c_2 :

$$y(1) = c_1 + c_2 = 2$$

$$y'(1) = 3c_1 - 2c_2 = 1$$

qui admet la solution

$$c_1 = 1, \quad c_2 = 1.$$

Donc, l'unique solution est

$$y(x) = x^3 + x^{-2}.$$

(b) **Résolution par Matlab symbolique.**—

```
dsolve('x^2*D2y=6*y', 'y(1)=2', 'Dy(1)=1', 'x')
y = (1+x^5)/x^2
```

(c) **Résolution par Matlab numérique.**— On récrit l'équation différentielle du second ordre en un système du premier ordre au moyen des variables

$$y(1) = y,$$

$$y(2) = y',$$

avec les conditions initiales en $x = 1$:

$$y_1(1) = 2, \quad y_2(1) = 1.$$

Alors,

$$y_1' = y_2,$$

$$y_2' = 6y_1/x^2.$$

Le fichier M `euler2.m` :

```
function yprime = euler2(x,y);
yprime = [y(2); 6*y(1)/x^2];
```

On appelle le solveur `ode23` et la commande `plot` :

```
xspan = [1 4]; % solution sur 1<=x<=4
y0 = [2; 1]; % conditions initiales
[x,y] = ode23('euler2',xspan,y0);
subplot(2,2,1); plot(x,y(:,1))
```

Le graphe de la solution numérique est à la fig. 2.6. □

EXEMPLE 2.10. Trouver la solution générale de l'équation d'Euler–Cauchy

$$x^2y'' + 7xy' + 9y = 0.$$

RÉSOLUTION. L'équation caractéristique

$$m^2 + 6m + 9 = (m + 3)^2 = 0$$

admet une racine double $m = -3$. La solution générale est donc

$$y(x) = (c_1 + c_2 \ln x)x^{-3}. \quad \square$$

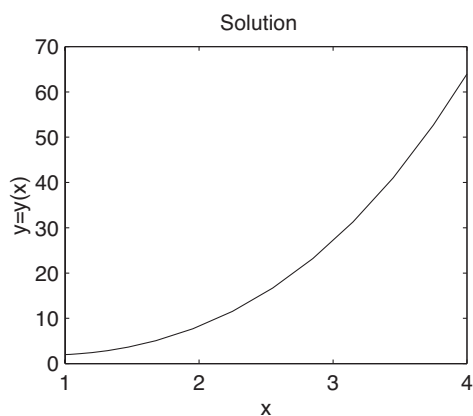


FIGURE 2.6. Graphe de la solution de l'équation linéaire de l'exemple 2.9.

EXEMPLE 2.11. Trouver la solution générale de l'équation d'Euler–Cauchy

$$x^2 y'' + 1.25y = 0.$$

RÉSOLUTION. L'équation caractéristique

$$m^2 - m + 1.25 = 0$$

admet deux racines complexes conjuguées l'une de l'autre

$$m_1 = \frac{1}{2} + i, \quad m_2 = \frac{1}{2} - i.$$

La solution générale est donc

$$y(x) = x^{1/2} [c_1 \cos(\ln x) + c_2 \sin(\ln x)]. \quad \square$$

On traitera de l'existence et de l'unicité de la solution des problèmes aux valeurs initiales au chapitre suivant.

Équations différentielles linéaires d'ordre quelconque

3.1. Équations homogènes

Considérons l'équation différentielle linéaire *non homogène* d'ordre n ,

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = r(x), \quad (3.1)$$

à coefficients variables, $a_0(x), a_1(x), \dots, a_{n-1}(x)$. Notons L l'opérateur différentiel du 1er membre :

$$L := D^n + a_{n-1}(x)D^{n-1} + \cdots + a_1(x)D + a_0(x)I, \quad D := ' = \frac{d}{dx}. \quad (3.2)$$

Alors, l'équation non homogène (3.1) s'écrit

$$Lu = r(x).$$

Si $r \equiv 0$, l'équation (3.1) est dite *homogène* :

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = 0, \quad (3.3)$$

c'est-à-dire

$$Ly = 0.$$

DÉFINITION 3.1. Une *solution* de (3.1) ou (3.3) sur $]a, b[$ est une fonction $y(x)$ n fois continûment dérivable sur $]a, b[$ qui satisfait identiquement l'équation différentielle.

Le théorème 2.1 se généralise aux équations linéaires homogènes d'ordre n quelconque.

THÉORÈME 3.1. Les **solutions** de (3.3) forment un espace vectoriel.

DÉMONSTRATION. Soit y_1, y_2, \dots, y_k , k solutions de $Ly = 0$. La linéarité de L implique :

$$L(c_1y_1 + c_2y_2 + \cdots + c_ky_k) = c_1Ly_1 + c_2Ly_2 + \cdots + c_kLy_k = 0, \quad c_i \in \mathbb{R}. \quad \square$$

DÉFINITION 3.2. On dit que n **fonctions**, f_1, f_2, \dots, f_n , sont *linéairement dépendantes* sur $]a, b[$ si, et seulement si, il existe n constantes non toutes nulles,

$$(k_1, k_2, \dots, k_n) \neq (0, 0, \dots, 0),$$

telles que

$$k_1f_1(x) + k_2f_2(x) + \cdots + k_nf_n(x) = 0, \quad \text{pour tout } x \in]a, b[. \quad (3.4)$$

Sinon, elles sont *linéairement indépendantes*.

REMARQUE 3.1. Soit f_1, f_2, \dots, f_n , n **fonctions** linéairement dépendantes. Sans perte de généralité, on peut supposer que $k_1 \neq 0$ dans (3.4). Alors f_1 est une combinaison linéaire de f_2, f_3, \dots, f_n .

$$f_1(x) = -\frac{1}{k_1} [k_2 f_2(x) + \dots + k_n f_n(x)].$$

On a le théorème d'existence et d'unicité suivant.

THÉORÈME 3.2. *Si les fonctions $a_0(x), a_1(x), \dots, a_{n-1}(x)$ sont continues sur $]a, b[$ et $x_0 \in]a, b[$, alors le problème aux valeurs initiales*

$$Ly = 0, \quad y(x_0) = k_1, \quad y'(x_0) = k_2, \quad \dots, \quad y^{(n-1)}(x_0) = k_n, \quad (3.5)$$

admet une et une seule solution.

DÉMONSTRATION. On peut démontrer le théorème en réduisant l'équation différentielle (3.3) d'ordre n à un système de n équations différentielles du 1er ordre. En effet, posons

$$u_1 = y, \quad u_2 = y', \quad \dots, \quad u_n = y^{(n-1)}.$$

Alors le problème aux valeurs initiales devient

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{bmatrix}' = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{bmatrix}, \quad \begin{bmatrix} u_1(x_0) \\ u_2(x_0) \\ \vdots \\ u_{n-1}(x_0) \\ u_n(x_0) \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_{n-1} \\ k_n \end{bmatrix}.$$

Sous forme matricielle, on a

$$\mathbf{u}'(x) = A(x)\mathbf{u}(x), \quad \mathbf{u}(x_0) = \mathbf{k}.$$

On dit que la matrice A est une matrice compagnon parce que le déterminant $|A - \lambda I|$ est le polynôme caractéristique de l'équation différentielle homogène :

$$|A - \lambda I| = (-1)^n (\lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0) = (-1)^n p_n(\lambda).$$

On peut démontrer par la méthode de Picard que ce système admet une et une seule solution. La récurrence de Picard est de la forme :

$$\mathbf{u}^{[n]}(x) = \mathbf{u}^{[0]}(x_0) + \int_{x_0}^x A(t)\mathbf{u}^{[n-1]}(t) dt, \quad \mathbf{u}^{[0]}(x_0) = \mathbf{k}. \quad \square$$

DÉFINITION 3.3. Le wronskien de n **fonctions**, $f_1(x), f_2(x), \dots, f_n(x)$, $n-1$ fois différentiables sur $]a, b[$ est le déterminant d'ordre n :

$$W(f_1, f_2, \dots, f_n)(x) := \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}. \quad (3.6)$$

On caractérise la dépendance linéaire de n **solutions** de l'équation linéaire homogène (3.3) au moyen de leur wronskien.

Montrons d'abord un lemme d'Abel.

LEMME 3.1 (Abel). Soit n **solutions**, y_1, y_2, \dots, y_n , de (3.3) sur $]a, b[$. Alors le wronskien $W(x) = W(y_1, y_2, \dots, y_n)(x)$ satisfait l'identité suivante :

$$W(x) = W(x_0) e^{-\int_{x_0}^x a_{n-1}(x) dx}, \quad x_0 \in]a, b[. \quad (3.7)$$

DÉMONSTRATION. Pour simplifier l'écriture, prenons $n = 3$; le cas général se traite de la même façon. Soit $W(x)$ le wronskien de trois solutions y_1, y_2, y_3 . La dérivée du wronskien est de la forme:

$$\begin{aligned} W'(x) &= \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{vmatrix}' \\ &= \begin{vmatrix} y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \\ y_1'' & y_2'' & y_3'' \end{vmatrix} + \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1'' & y_2'' & y_3'' \\ y_1'' & y_2'' & y_3'' \end{vmatrix} + \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1''' & y_2''' & y_3''' \end{vmatrix} \\ &= \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1''' & y_2''' & y_3''' \end{vmatrix} \\ &= \begin{vmatrix} & y_1 & & & y_2 & & & y_3 \\ & y_1' & & & y_2' & & & y_3' \\ -a_0 y_1 - a_1 y_1' - a_2 y_1'' & & & -a_0 y_2 - a_1 y_2' - a_2 y_2'' & & & -a_0 y_3 - a_1 y_3' - a_2 y_3'' & \end{vmatrix}, \end{aligned}$$

puisque les deux premiers déterminants sont nuls et dans le dernier déterminant on emploie le fait que y_k , $k = 1, 2, 3$, est une solution de l'équation homogène (3.3).

Si l'on additionne a_0 fois la 1ère ligne et a_1 fois la 2ème ligne à la 3ème ligne, on obtient

$$W'(x) = -a_2(x)W(x).$$

C'est une équation différentielle séparable:

$$\frac{dW}{W} = -a_2(x) dx.$$

La solution est

$$\ln |W| = -\int a_2(x) dx + c,$$

c'est-à-dire

$$W(x) = W(x_0) e^{-\int_{x_0}^x a_2(x) dx}, \quad x_0 \in]a, b[. \quad \square$$

THÉORÈME 3.3. Si les coefficients $a_0(x), a_1(x), \dots, a_{n-1}(x)$ de (3.3) sont continus sur $]a, b[$, alors n **solutions**, y_1, y_2, \dots, y_n , de (3.3) sont linéairement dépendantes si et seulement si leur wronskien s'annule en un point quelconque $x_0 \in]a, b[$:

$$W(y_1, y_2, \dots, y_n)(x_0) := \begin{vmatrix} y_1(x_0) & \cdots & y_n(x_0) \\ y_1'(x_0) & \cdots & y_n'(x_0) \\ \vdots & & \vdots \\ y_1^{(n-1)}(x_0) & \cdots & y_n^{(n-1)}(x_0) \end{vmatrix} = 0. \quad (3.8)$$

DÉMONSTRATION. Si les solutions sont linéairement dépendantes, alors par la définition 3.2 il existe n constantes non toutes nulles,

$$(k_1, k_2, \dots, k_n) \neq (0, 0, \dots, 0),$$

telles que

$$k_1 y_1(x) + k_2 y_2(x) + \cdots + k_n y_n(x) = 0, \quad \text{pour tout } x \in]a, b[.$$

Si l'on dérive cette identité $n - 1$ fois, on obtient

$$\begin{aligned} k_1 y_1(x) + k_2 y_2(x) + \cdots + k_n y_n(x) &= 0, \\ k_1 y_1'(x) + k_2 y_2'(x) + \cdots + k_n y_n'(x) &= 0, \\ &\vdots \\ k_1 y_1^{(n-1)}(x) + k_2 y_2^{(n-1)}(x) + \cdots + k_n y_n^{(n-1)}(x) &= 0. \end{aligned}$$

Ce système linéaire homogène en k_1, k_2, \dots, k_n se réécrit sous la forme matricielle :

$$\begin{bmatrix} y_1(x) & \cdots & y_n(x) \\ y_1'(x) & \cdots & y_n'(x) \\ \vdots & & \vdots \\ y_1^{(n-1)}(x) & \cdots & y_n^{(n-1)}(x) \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (3.9)$$

c'est-à-dire

$$A\mathbf{k} = 0.$$

Puisque, par hypothèse, la solution \mathbf{k} est non nulle, le déterminant du système doit être nul,

$$\det A = W(y_1, y_2, \dots, y_n)(x) = 0, \quad \text{pour tout } x \in]a, b[.$$

D'autre part, si le wronskien de n **solutions** s'annule en un point $x_0 \in]a, b[$,

$$W(y_1, y_2, \dots, y_n)(x_0) = 0,$$

il s'annule pour tout $x \in]a, b[$ par le lemme 3.1 d'Abel. Alors le déterminant $W(x)$ du système (3.9) est nul pour tout $x \in]a, b[$. Donc ce système admet une solution \mathbf{k} non nulle. Par conséquent les solutions y_1, y_2, \dots, y_n de (3.3) sont linéairement dépendantes. \square

REMARQUE 3.2. Le wronskien de n fonctions linéairement dépendantes sur $]a, b[$ est nécessairement nul sur $]a, b[$, comme on peut voir à la 1ère partie de la démonstration du théorème 3.3. Mais pour des fonctions qui *ne sont pas solutions* d'une équation différentielle linéaire homogène, le fait que le wronskien soit nul sur $]a, b[$ n'est pas une condition suffisante pour que ces fonctions soient dépendantes. Par exemple, les fonctions $u_1 = x^3$ et $u_2 = |x|^3$ sont de classe $C^1[-1, 1]$ et sont linéairement indépendantes sur l'intervalle $[-1, 1]$, mais leur wronskien est identiquement nul sur cet intervalle.

COROLLAIRE 3.1. *Si les coefficients $a_0(x), a_1(x), \dots, a_{n-1}(x)$ de (3.3) sont continus sur $]a, b[$, alors n **solutions**, y_1, y_2, \dots, y_n , de (3.3) sont linéairement indépendantes si et seulement si leur wronskien ne s'annule pas en un seul point $x_0 \in]a, b[$.*

COROLLAIRE 3.2. *Soit $f_1(x), f_2(x), \dots, f_n(x)$ n fonctions de classe C^n sur un intervalle réel I . Si $W(f_1, \dots, f_n)(x) \neq 0$ sur I , alors il existe une unique équation différentielle homogène d'ordre n (avec le coefficient de $y^{(n)}$ l'unité) :*

$$(-1)^n \frac{W(y, f_1, \dots, f_n)}{W(f_1, \dots, f_n)} = 0,$$

pour laquelle ces fonctions forment un système de n solutions indépendantes.

EXEMPLE 3.1. Montrer que les fonctions

$$y_1(x) = \cosh x \quad \text{et} \quad y_2(x) = \sinh x$$

sont linéairement indépendantes.

RÉSOLUTION. Puisque y_1'' et y_2'' sont continues, par le corollaire 3.2, y_1 et y_2 sont solutions d'une même équation différentielle. Leur wronskien satisfait l'identité

$$W(y_1, y_2)(x) = \begin{vmatrix} \cosh x & \sinh x \\ \sinh x & \cosh x \end{vmatrix} = \cosh^2 x - \sinh^2 x = 1,$$

pour tout x . Donc y_1 et y_2 sont indépendantes. On voit facilement que y_1 et y_2 sont solutions de l'équation différentielle

$$y'' - y = 0.$$

□

Dans la résolution on a employé l'identité suivante :

$$\begin{aligned} \cosh^2 x - \sinh^2 x &= \left(\frac{e^x + e^{-x}}{2} \right)^2 - \left(\frac{e^x - e^{-x}}{2} \right)^2 \\ &= \frac{1}{4} (e^{2x} + e^{-2x} + 2e^x e^{-x} - e^{2x} - e^{-2x} + 2e^x e^{-x}) \\ &= 1. \end{aligned}$$

EXEMPLE 3.2. Utiliser le wronskien pour montrer que les fonctions

$$y_1(x) = x^m \quad \text{et} \quad y_2(x) = x^m \ln x$$

sont linéairement indépendantes sur $x > 0$ et construire une équation différentielle du second ordre pour laquelle ces fonctions sont solutions.

RÉSOLUTION. On vérifie que le wronskien de y_1 et y_2 ne s'annule pas sur $x > 0$:

$$\begin{aligned} W(y_1, y_2)(x) &= \begin{vmatrix} x^m & x^m \ln x \\ mx^{m-1} & mx^{m-1} \ln x + x^{m-1} \end{vmatrix} \\ &= x^m x^{m-1} \begin{vmatrix} 1 & \ln x \\ m & m \ln x + 1 \end{vmatrix} \\ &= x^{2m-1} (1 + m \ln x - m \ln x) = x^{2m-1} \neq 0, \quad \text{pour tout } x > 0. \end{aligned}$$

Alors, par le corollaire 3.2, y_1 et y_2 sont linéairement indépendantes. Par le même corollaire,

$$\begin{aligned} W(y, x^m, x^m \ln x)(x) &= \begin{vmatrix} y & x^m & x^m \ln x \\ y' & mx^{m-1} & mx^{m-1} \ln x + x^{m-1} \\ y'' & m(m-1)x^{m-2} & m(m-1)x^{m-2} \ln x + (2m-1)x^{m-2} \end{vmatrix} \\ &= 0. \end{aligned}$$

Pour évaluer ce déterminant, on multiplie la 2e et la 3e lignes respectivement par x et x^2 , on divise la 2e et la 3e colonnes par x^m , on soustrait m fois la 1ère ligne

de la seconde ligne et $m(m-1)$ fois la 1ère ligne de la 3e ligne :

$$\begin{vmatrix} y & 1 & \ln x \\ xy' - my & 0 & 1 \\ x^2y'' - m(m-1)y & 0 & 2m-1 \end{vmatrix} = 0.$$

Enfin, on développe ce déterminant suivant la 2e colonne, ce qui donne l'équation d'Euler-Cauchy :

$$x^2y'' + (1-2m)xy' + m^2y = 0. \quad \square$$

DÉFINITION 3.4. On appelle *système fondamental* ou *base* sur $]a, b[$ n solutions, y_1, y_2, \dots, y_n , de l'équation homogène (3.3) linéairement indépendantes sur $]a, b[$.

DÉFINITION 3.5. Soit y_1, y_2, \dots, y_n un système fondamental pour (3.3). On appelle *solution générale* de (3.3) sur $]a, b[$ une solution de la forme

$$y(x) = c_1y_1(x) + c_2y_2(x) + \dots + c_ny_n(x), \quad (3.10)$$

où c_1, c_2, \dots, c_n sont n constantes arbitraires.

THÉORÈME 3.4. Si les fonctions $a_0(x), a_1(x), \dots, a_{n-1}(x)$ sont continues sur $]a, b[$, alors l'équation linéaire homogène (3.3) admet une solution générale sur $]a, b[$.

DÉMONSTRATION. Par le théorème 3.2, pour $i = 1, 2, \dots, n$, le problème aux valeurs initiales (3.5),

$$Ly = 0, \quad k_i = 1, \quad k_j = 0 \quad j \neq i,$$

admet une (et une seule) solution $y_i(x)$ telle que

$$y_i^{(i-1)}(x_0) = 1, \quad y_i^{(j-1)}(x_0) = 0, \quad j = 1, 2, \dots, i-1, i+1, \dots, n.$$

Alors le wronskien

$$W(y_1, y_2, \dots, y_n)(x_0) = \begin{vmatrix} y_1(x_0) & \dots & y_n(x_0) \\ y_1'(x_0) & \dots & y_n'(x_0) \\ \vdots & & \vdots \\ y_1^{(n-1)}(x_0) & \dots & y_n^{(n-1)}(x_0) \end{vmatrix} = |I_n| = 1,$$

où I_n est la matrice identité d'ordre n . Les solutions sont donc indépendantes par le corollaire 3.1. \square

THÉORÈME 3.5. Si les fonctions $a_0(x), a_1(x), \dots, a_{n-1}(x)$ sont continues sur $]a, b[$, alors toute solution du problème aux valeurs initiales (3.5) sur $]a, b[$ s'obtient au moyen d'une solution générale.

DÉMONSTRATION. Soit

$$y = c_1y_1 + c_2y_2 + \dots + c_ny_n$$

une solution générale de (3.3). Le système

$$\begin{bmatrix} y_1(x_0) & \dots & y_n(x_0) \\ y_1'(x_0) & \dots & y_n'(x_0) \\ \vdots & & \vdots \\ y_1^{(n-1)}(x_0) & \dots & y_n^{(n-1)}(x_0) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$$

admet une solution unique \mathbf{c} puisque le déterminant du système est non nul. \square

On peut trouver, par variation des paramètres, une seconde solution d'une équation du second ordre à coefficients variables si l'on en connaît une première solution.

EXEMPLE 3.3. Soit

$$y_1(x) = e^{x^2}$$

une solution de l'équation différentielle du 2ème ordre :

$$Ly := y'' - 4xy' + (4x^2 - 2)y = 0.$$

Trouver une seconde solution indépendante.

RÉSOLUTION. (a) **Résolution analytique.**— On utilise la variation des paramètres. En posant

$$y_2(x) = u(x)y_1(x),$$

on obtient

$$\begin{aligned} (4x^2 - 2)y_2 &= (4x^2 - 2)uy_1, \\ -4xy_2' &= -4xuy_1' - 4xu'y_1, \\ y_2'' &= uy_1'' + 2u'y_1' + u''y_1. \end{aligned}$$

On additionne respectivement les 1ers et les 2èmes membres :

$$Ly_2 = uLy_1 + (2y_1' - 4xy_1)u' + y_1u''.$$

On a $Ly_1 = 0$ et $Ly_2 = 0$ puisque y_1 est une solution et l'on veut que y_2 soit une solution. Maintenant on remplace y_1 par son expression dans l'équation différentielle en u :

$$e^{x^2}u'' + (4xe^{x^2} - 4xe^{x^2})u' = 0,$$

d'où

$$u'' = 0 \implies u' = k_1 \implies u = k_1x + k_2$$

et

$$y_2(x) = (k_1x + k_2)e^{x^2}.$$

Il suffit de prendre $k_2 = 0$ parce que $k_2e^{x^2}$ est contenu dans y_1 et $k_1 = 1$ parce que xe^{x^2} sera multiplié par une constante arbitraire. Alors la solution générale est de la forme

$$y(x) = (c_1 + c_2x)e^{x^2}.$$

(b) **Résolution par Matlab symbolique.**—

```
dsolve('D2y-4*x*Dy+(4*x^2-2)*y=0','x')
y = C1*exp(x^2)+C2*exp(x^2)*x
```

□

3.2. Équations linéaires homogènes à coefficients constants

Considérons l'équation différentielle linéaire homogène d'ordre n ,

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0, \quad (3.11)$$

à coefficients constants, a_0, a_1, \dots, a_{n-1} . Notons L l'opérateur différentiel du 1er membre :

$$L := D^n + a_{n-1}D^{n-1} + \cdots + a_1D + a_0I, \quad D := ' = \frac{d}{dx}. \quad (3.12)$$

On pose $y(x) = e^{\lambda x}$ dans (3.11). On obtient alors l'équation caractéristique

$$p(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0, \quad (3.13)$$

Si les n racines de $p(\lambda) = 0$ sont distinctes, on a n solutions indépendantes :

$$y_1(x) = e^{\lambda_1 x}, \quad y_2(x) = e^{\lambda_2 x}, \quad \dots, \quad y_n(x) = e^{\lambda_n x}, \quad (3.14)$$

et la solution générale est de la forme

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} + \cdots + c_n e^{\lambda_n x}. \quad (3.15)$$

Si (3.13) admet une racine double, disons, $\lambda_1 = \lambda_2$, on a deux solutions indépendantes de la forme

$$y_1(x) = e^{\lambda_1 x}, \quad y_2(x) = x e^{\lambda_1 x}.$$

De même, si l'on a une racine triple, disons, $\lambda_1 = \lambda_2 = \lambda_3$, on a trois solutions indépendantes de la forme

$$y_1(x) = e^{\lambda_1 x}, \quad y_2(x) = x e^{\lambda_1 x}, \quad y_3(x) = x^2 e^{\lambda_1 x}.$$

On démontre le théorème suivant.

THÉORÈME 3.6. *Soit μ une racine de multiplicité m de l'équation caractéristique (3.13). Alors l'équation différentielle (3.11) admet m solutions indépendantes de la forme*

$$y_1(x) = e^{\mu x}, \quad y_2(x) = x e^{\mu x}, \quad \dots, \quad y_m(x) = x^{m-1} e^{\mu x}. \quad (3.16)$$

DÉMONSTRATION. Écrivons

$$p(D)y = (D^n + a_{n-1}D^{n-1} + \cdots + a_1D + a_0)y = 0.$$

Puisque, par hypothèse,

$$p(\lambda) = q(\lambda)(\lambda - \mu)^m,$$

et que les coefficients sont constants, l'opérateur différentiel s'écrit sous la forme

$$p(D) = q(D)(D - \mu)^m.$$

On voit par récurrence que les m fonctions (3.16),

$$x^k e^{\mu x}, \quad k = 0, 1, \dots, m-1,$$

satisfont les équations suivantes :

$$\begin{aligned} (D - \mu) (x^k e^{\mu x}) &= kx^{k-1} e^{\mu x} + \mu x^k e^{\mu x} - \mu x^k e^{\mu x} \\ &= kx^{k-1} e^{\mu x}, \\ (D - \mu)^2 (x^k e^{\mu x}) &= (D - \mu) (kx^{k-1} e^{\mu x}) \\ &= k(k-1)x^{k-2} e^{\mu x}, \\ &\vdots \\ (D - \mu)^k (x^k e^{\mu x}) &= k! e^{\mu x}, \\ (D - \mu)^{k+1} (x^k e^{\mu x}) &= k! (\mu e^{\mu x} - \mu e^{\mu x}) = 0. \end{aligned}$$

Puisque $m \geq k + 1$, on a

$$(D - \mu)^m (x^k e^{\mu x}) = 0, \quad k = 0, 1, \dots, m - 1.$$

Donc, par le lemme 3.2 qui suit, les fonctions (3.16) forment m solutions indépendantes de (3.11). \square

LEMME 3.2. *Soit*

$$y_1(x) = e^{\mu x}, \quad y_2(x) = x e^{\mu x}, \quad \dots, \quad y_m(x) = x^{m-1} e^{\mu x},$$

m solutions d'une équation différentielle linéaire homogène. Alors elles sont indépendantes.

DÉMONSTRATION. Par le corollaire 3.1, il suffit de montrer que le wronskien des solutions ne s'annule pas en $x = 0$. On a vu dans la démonstration du théorème précédent que

$$(D - \mu)^k (x^k e^{\mu x}) = k! e^{\mu x},$$

c'est-à-dire

$$D^k (x^k e^{\mu x}) = k! e^{\mu x} + \text{des termes en } x^l e^{\mu x}, \quad l = 1, 2, \dots, k - 1.$$

Donc

$$D^k (x^k e^{\mu x}) \Big|_{x=0} = k!, \quad D^k (x^{k+l} e^{\mu x}) \Big|_{x=0} = 0, \quad l \geq 1.$$

Il suit que la matrice M du wronskien est triangulaire inférieure avec $m_{i,i} = (i - 1)!$,

$$W(0) = \begin{bmatrix} 0! & 0 & 0 & \dots & 0 \\ \times & 1! & 0 & & 0 \\ \times & \times & 2! & 0 & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \times & \times & \dots & \times & (m-1)! \end{bmatrix} \neq 0. \quad \square$$

EXEMPLE 3.4. Trouver la solution générale de

$$(D^4 - 13D^2 + 36I)y = 0.$$

RÉSOLUTION. Le polynôme caractéristique se met facilement en facteur :

$$\begin{aligned} \lambda^4 - 13\lambda^2 + 36 &= (\lambda^2 - 9)(\lambda^2 - 4) \\ &= (\lambda + 3)(\lambda - 3)(\lambda + 2)(\lambda - 2). \end{aligned}$$

Alors :

$$y(x) = c_1 e^{-3x} + c_2 e^{3x} + c_3 e^{-2x} + c_4 e^{2x}.$$

Le solveur de polynôme de Matlab.— Pour trouver les zéros du polynôme caractéristique

$$\lambda^4 - 13\lambda^2 + 36$$

avec Matlab, on représente le polynôme par le vecteur de ses coefficients,

$$p = [1 \quad 0 \quad -13 \quad 0 \quad 36]$$

et l'on applique la commande `roots` sur p .

```
>> p = [1 0 -13 0 36]
p = 1    0   -13    0    36
>> r = roots(p)
r =
  3.0000
 -3.0000
  2.0000
 -2.0000
```

De fait, la commande `roots` construit une matrice compagnon C de p (voir la démonstration du théorème 3.2)

$$C = \begin{bmatrix} 0 & 13 & 0 & -36 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

et emploie l'algorithme QR pour trouver les valeurs propres de C qui, de fait, sont les zéros de p .

```
>> p = [1 0 -13 0 36];
>> C = compan(p)
C =
  0    13    0   -36
  1     0    0    0
  0     1    0    0
  0     0    1    0
>> eigenvalues = eig(C)'
eigenvalues = 3.0000   -3.0000    2.0000   -2.0000
```

□

EXEMPLE 3.5. Trouver la solution générale de

$$(D - I)^3 y = 0.$$

RÉSOLUTION. Le polynôme caractéristique $(\lambda - 1)^3$ admet un zéro triple :

$$\lambda_1 = \lambda_2 = \lambda_3 = 1,$$

Alors :

$$y(x) = c_1 e^x + c_2 x e^x + c_3 x^2 e^x. \quad \square$$

Si l'équation caractéristique (3.13) admet deux racines complexes, $\lambda_{1,2} = \alpha \pm i\beta$, on a deux solutions indépendantes :

$$y_1(x) = e^{\alpha x} \cos \beta x, \quad y_2(x) = e^{\alpha x} \sin \beta x,$$

comme au chapitre 2.

Si l'ordre de l'équation différentielle est au moins 4, $n \geq 4$, il y a d'autres possibilités pour les racines complexes. Si l'on a deux paires distinctes de racines conjuguées complexes, par exemple, $\lambda_{1,2} = \alpha_1 \pm i\beta_1$ et $\lambda_{3,4} = \alpha_2 \pm i\beta_2$ où $(\alpha_1, \beta_1) \neq (\alpha_2, \beta_2)$, alors on a les solutions indépendantes

$$\begin{aligned} y_1(x) &= e^{\alpha_1 x} \cos \beta_1 x, & y_2(x) &= e^{\alpha_1 x} \sin \beta_1 x, \\ y_3(x) &= e^{\alpha_2 x} \cos \beta_2 x, & y_4(x) &= e^{\alpha_2 x} \sin \beta_2 x. \end{aligned}$$

Mais si l'on a des racines complexes multiples, par exemple, $\lambda_1 = \lambda_2 = \alpha + i\beta$ et $\lambda_3 = \lambda_4 = \alpha - i\beta$, alors on a les solutions indépendantes suivantes :

$$\begin{aligned} y_1(x) &= e^{\alpha x} \cos \beta x, & y_2(x) &= e^{\alpha x} \sin \beta x, \\ y_3(x) &= x e^{\alpha x} \cos \beta x, & y_4(x) &= x e^{\alpha x} \sin \beta x, \end{aligned}$$

qu'on peut écrire dans un ordre arbitraire.

EXEMPLE 3.6. Trouver la solution générale de l'équation différentielle

$$y^{(4)} + 13y'' + 36y = 0.$$

SOLUTION. L'équation caractéristique

$$\lambda^4 + 13\lambda^2 + 36 = (\lambda^2 + 4)(\lambda^2 + 9) = 0$$

admet les racines $\lambda_{1,2} = \pm 2i$ et $\lambda_{3,4} = \pm 3i$. Donc, la solution générale est

$$y(x) = c_1 \cos 2x + c_2 \sin 2x + c_3 \cos 3x + c_4 \sin 3x. \quad \square$$

EXEMPLE 3.7. Trouver la solution générale de l'équation différentielle

$$y^{(4)} + 8y''' + 26y'' + 40y' + 25y = 0.$$

SOLUTION. L'équation caractéristique

$$\lambda^4 + 8\lambda^3 + 26\lambda^2 + 40\lambda + 25 = (\lambda^2 + 4\lambda + 5)^2 = 0$$

admet les racines $\lambda_{1,2} = -2 + i$ et $\lambda_{3,4} = -2 - i$. Donc, la solution générale est

$$y(x) = c_1 e^{-2x} \cos x + c_2 e^{-2x} \sin x + c_3 x e^{-2x} \cos x + c_4 x e^{-2x} \sin x. \quad \square$$

On résout les équations d'Euler–Cauchy homogènes d'ordre supérieur à 2 comme au chapitre 2.

EXEMPLE 3.8. Trouver la solution générale de l'équation d'Euler–Cauchy

$$x^3 y''' - 3x^2 y'' + 6xy' - 6y = 0.$$

RÉSOLUTION. (a) Résolution analytique.— En posant

$$y = x^m$$

dans l'équation différentielle, on obtient

$$m(m-1)(m-2)x^m - 3m(m-1)x^m + 6mx^m - 6x^m = 0,$$

d'où l'équation caractéristique :

$$m(m-1)(m-2) - 3m(m-1) + 6m - 6 = 0.$$

Si l'on remarque que $m-1$ est un facteur commun, on obtient

$$\begin{aligned} (m-1)[m(m-2) - 3m + 6] &= (m-1)(m^2 - 5m + 6) \\ &= (m-1)(m-2)(m-3) = 0. \end{aligned}$$

Alors :

$$y(x) = c_1 x + c_2 x^2 + c_3 x^3.$$

(b) Résolution par Matlab symbolique.—

```
dsolve('x^3*D3y-3*x^2*D2y+6*x*Dy-6*y=0', 'x')
y = C1*x+C2*x^2+C3*x^3
```

□

3.3. Équations linéaires non homogènes

Considérons l'équation différentielle linéaire *non homogène* d'ordre n ,

$$Ly := y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = r(x). \quad (3.17)$$

Soit

$$y_h(x) = c_1 y_1(x) + c_2 y_2(x) + \dots + c_n y_n(x), \quad (3.18)$$

une solution générale de l'équation homogène

$$Ly = 0.$$

De plus, soit $y_p(x)$ une *solution particulière* de l'équation non homogène (3.17). Alors,

$$y_g(x) = y_h(x) + y_p(x)$$

est la solution générale de (3.17). En effet,

$$Ly_g = Ly_h + Ly_p = 0 + r(x).$$

EXEMPLE 3.9. Trouver la solution générale $y_g(x)$ de

$$y'' - y = 3e^{2x}$$

si

$$y_p(x) = e^{2x}$$

est une solution particulière.

RÉSOLUTION. **(a) Résolution analytique.**— Il est facile de voir que e^{2x} est bien une solution particulière. Puisque

$$y'' - y = 0 \implies \lambda^2 - 1 = 0 \implies \lambda = \pm 1,$$

alors

$$y_h(x) = c_1 e^x + c_2 e^{-x}$$

et

$$y_g(x) = c_1 e^x + c_2 e^{-x} + e^{2x}.$$

(b) Résolution par Matlab symbolique.—

```
dsolve('D2y-y-3*exp(2*x)', 'x')
y = (exp(2*x)*exp(x)+C1*exp(x)^2+C2)/exp(x)
z = expand(y)
z = exp(x)^2+exp(x)*C1+1/exp(x)*C2
```

□

Voici une seconde méthode de résolution d'une équation linéaire non homogène du 1er ordre,

EXEMPLE 3.10. Trouver la solution générale de l'équation linéaire non homogène du 1er ordre,

$$Ly := y' + f(x)y = r(x). \quad (3.19)$$

RÉSOLUTION. L'équation homogène $Ly = 0$ est séparable :

$$\frac{dy}{y} = -f(x) dx \implies \ln|y| = -\int f(x) dx \implies y_h(x) = e^{-\int f(x) dx}.$$

On trouve une solution particulière par variation des paramètres en posant

$$y_p(x) = u(x)y_h(x)$$

dans l'équation non homogène $Ly = r(x)$:

$$\begin{aligned} y_p' &= uy_h' + u'y_h \\ f(x)y_p &= uf(x)y_h. \end{aligned}$$

On additionne chacun des deux membres de ces deux expressions :

$$Ly_p = uLy_h + u'y_h = u'y_h = r(x).$$

L'équation différentielle $u'y_h = r$ est séparable :

$$du = e^{\int f(x) dx} r(x) dx.$$

Alors

$$u(x) = \int e^{\int f(x) dx} r(x) dx.$$

Donc la solution générale $y = cy_h + y_p$ de (3.19) est :

$$y(x) = ce^{-\int f(x) dx} + e^{-\int f(x) dx} \int e^{\int f(x) dx} r(x) dx. \quad \square$$

Dans les deux sections qui suivent, on présente deux méthodes pour trouver une solution particulière, soit la méthode des coefficients indéterminés et la méthode de la variation des paramètres. La première méthode, plus particulière que la seconde, ne requiert pas toujours la solution générale de l'équation homogène, à la différence de la deuxième méthode.

3.4. Méthode des coefficients indéterminés

Considérons l'équation différentielle linéaire non homogène d'ordre n ,

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = r(x), \quad (3.20)$$

à coefficients constants, a_0, a_1, \dots, a_{n-1} .

Si l'espace des dérivées du 2ème membre de (3.20) est de dimension finie, on peut employer la méthode des coefficients indéterminés.

Voici des exemples de fonctions $r(x)$ qui admettent un nombre fini de dérivées linéairement indépendantes; de plus, on indique la dimension de l'espace des

dérivées.

$$\begin{aligned} r(x) &= x^2 + 2x + 1, & r'(x) &= 2x + 2, & r''(x) &= 2, \\ r^{(k)}(x) &= 0, \quad k = 3, 4, \dots, & & \implies \dim. = 3; \\ r(x) &= \cos 2x + \sin 2x, & r'(x) &= -2 \sin 2x + 2 \cos 2x, \\ r''(x) &= -4r(x), & & \implies \dim. = 2; \\ r(x) &= x e^x, & r'(x) &= e^x + x e^x, \\ r''(x) &= 2r'(x) - r(x), & & \implies \dim. = 2. \end{aligned}$$

Parmi les fonctions qui admettent un nombre fini de dérivées indépendantes, on a les polynômes, les exponentielles, sinus, cosinus, sinus hyperbolique, cosinus hyperbolique, leurs sommes et leurs produits.

La méthode des coefficients indéterminés consiste à prendre une solution particulière qui est une combinaison linéaire,

$$y_p(x) = c_1 p_1(x) + c_2 p_2(x) + \dots + c_l p_l(x),$$

des dérivées indépendantes de la fonction $r(x)$. On détermine les coefficients c_k en substituant $y_p(x)$ dans (3.20). Si l'on obtient une contradiction, on a fait un mauvais choix ou une erreur.

EXEMPLE 3.11. Trouver la solution générale $y_g(x)$ de

$$Ly := y'' + y = 3x^2$$

par la méthode des coefficients indéterminés.

RÉSOLUTION. (a) **Résolution analytique.**— Posons

$$y_p(x) = ax^2 + bx + c$$

dans l'équation différentielle et additionnons chacun des deux membres :

$$\begin{aligned} y_p &= ax^2 + bx + c \\ y_p'' &= 2a \\ Ly_p &= ax^2 + bx + (2a + c) \\ &= 3x^2. \end{aligned}$$

En identifiant les coefficients de 1, x et x^2 des deux derniers membres, on obtient :

$$a = 3, \quad b = 0, \quad c = -2a = -6.$$

La solution générale de $Ly = 0$ est

$$y_h(x) = A \cos x + B \sin x.$$

Alors la solution générale de $Ly = 3x^2$ est

$$y_g(x) = A \cos x + B \sin x + 3x^2 - 6.$$

(b) **Résolution par Matlab symbolique.**—

```
dsolve('D2y+y=3*x^2','x')
y = -6+3*x^2+C1*sin(x)+C2*cos(x)
```

□

Remarque importante. Si pour un terme choisi $p_j(x)$, $x^k p_j(x)$ est une solution de l'équation homogène, mais pas $x^{k+1} p_j(x)$, alors il faut remplacer $p_j(x)$ par $x^{k+1} p_j(x)$.

EXEMPLE 3.12. Trouver la forme d'une solution particulière de l'équation

$$y'' - 4y' + 4y = 3e^{2x} + 32 \sin x$$

par coefficients indéterminés.

RÉSOLUTION. Puisque la solution générale de l'équation homogène est

$$y_h(x) = c_1 e^{2x} + c_2 x e^{2x},$$

une solution particulière est de la forme

$$y_p(x) = ax^2 e^{2x} + b \cos x + c \sin x. \quad \square$$

EXEMPLE 3.13. Résoudre

$$y''' - y' = 4e^{-x} + 3e^{2x}, \quad y(0) = 0, \quad y'(0) = -1, \quad y''(0) = 2$$

et tracer la solution.

RÉSOLUTION. **(a) Résolution analytique.**— L'équation caractéristique est

$$\lambda^3 - \lambda = \lambda(\lambda - 1)(\lambda + 1) = 0.$$

La solution générale de l'équation homogène est

$$y_h(x) = c_1 + c_2 e^x + c_3 e^{-x}.$$

Puisque e^{-x} apparaît au second membre de l'équation différentielle, on choisit une solution particulière de la forme

$$y_p(x) = ax e^{-x} + b e^{2x}.$$

Alors,

$$\begin{aligned} y_p'(x) &= -ax e^{-x} + a e^{-x} + 2b e^{2x}, \\ y_p''(x) &= ax e^{-x} - 2a e^{-x} + 4b e^{2x}, \\ y_p'''(x) &= -ax e^{-x} + 3a e^{-x} + 8b e^{2x}. \end{aligned}$$

Donc,

$$\begin{aligned} y_p'''(x) - y_p'(x) &= 2a e^{-x} + 6b e^{2x} \\ &= 4e^{-x} + 3e^{2x}, \quad \text{pour tout } x. \end{aligned}$$

Si l'on identifie les coefficients de e^{-x} et e^{2x} , on a

$$a = 2, \quad b = \frac{1}{2},$$

ce qui donne la solution particulière de l'équation non homogène

$$y_p(x) = 2x e^{-x} + \frac{1}{2} e^{2x}$$

et la solution générale de l'équation non homogène

$$y(x) = c_1 + c_2 e^x + c_3 e^{-x} + 2x e^{-x} + \frac{1}{2} e^{2x}.$$

On détermine les constantes arbitraires c_1 , c_2 et c_3 au moyen des conditions initiales :

$$\begin{aligned}y(0) &= c_1 + c_2 + c_3 + \frac{1}{2} = 0, \\y'(0) &= c_2 - c_3 + 3 = -1, \\y''(0) &= c_2 + c_3 - 2 = 2,\end{aligned}$$

d'où le système algébrique linéaire

$$\begin{aligned}c_1 + c_2 + c_3 &= -\frac{1}{2}, \\c_2 - c_3 &= -4, \\c_2 + c_3 &= 4,\end{aligned}$$

qui admet la solution

$$c_1 = -\frac{9}{2}, \quad c_2 = 0, \quad c_3 = 4.$$

Donc l'unique solution :

$$y(x) = -\frac{9}{2} + 4e^{-x} + 2xe^{-x} + \frac{1}{2}e^{2x}.$$

(b) Résolution par Matlab symbolique.—

```
dsolve('D3y-Dy=4*exp(-x)+3*exp(2*x)', 'y(0)=0', 'Dy(0)=-1', 'D2y(0)=2', 'x')
y = 1/2*(8+exp(3*x)+4*x-9*exp(x))/exp(x)
z =expand(y)
z = 4/exp(x)+1/2*exp(x)^2+2/exp(x)*x-9/2
```

(c) Résolution par Matlab numérique.— On récrit l'équation différentielle du 3e ordre au moyen des variables

$$\begin{aligned}y(1) &= y, \\y(2) &= y', \\y(3) &= y''.\end{aligned}$$

Alors

$$\begin{aligned}y(1)' &= y(2), \\y(2)' &= y(3), \\y(3)' &= y(2) + 4 * \exp(-x) + 3 * \exp(2 * x).\end{aligned}$$

Le fichier M `exp39.m` :

```
function yprime = exp39(x,y);
yprime=[y(2); y(3); y(2)+4*exp(-x)+3*exp(2*x)];
```

On appelle le solveur `ode23` et la commande `plot` :

```
xspan = [0 2]; % solution pour 0<=x<=2
y0 = [0;-1;2]; % conditions initiales
[x,y] = ode23('exp39',xspan,y0);
plot(x,y(:,1))
```

Le graphe de la solution numérique est dans la fig. 3.1. □

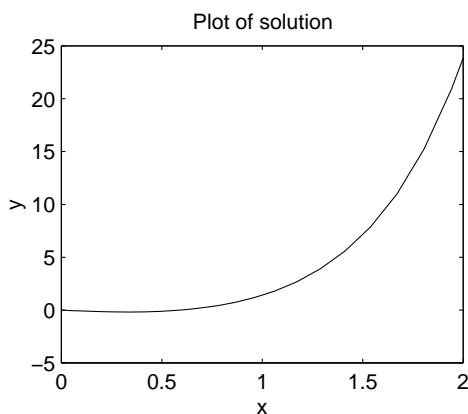


FIGURE 3.1. Graphe de la solution de l'équation linéaire de l'exemple 3.13.

3.5. Solution particulière par variation des paramètres

Considérons l'équation différentielle linéaire *non homogène* d'ordre n ,

$$Ly := y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = r(x), \quad (3.21)$$

sous forme standard, c'est-à-dire, le coefficient de $y^{(n)}$ est égal à 1.

Soit

$$y_h(x) = c_1y_1(x) + c_2y_2(x) + \cdots + c_ny_n(x), \quad (3.22)$$

une solution générale de l'équation homogène

$$Ly = 0.$$

Par simplicité, on dérive la méthode de la variation des paramètres dans le cas $n = 3$; le cas général suit de la même façon.

Suivant une idée de Lagrange, on prend une solution particulière de la forme

$$y_p(x) = c_1(x)y_1(x) + c_2(x)y_2(x) + c_3(x)y_3(x), \quad (3.23)$$

où l'on a fait varier les paramètres c_1 , c_2 et c_3 de la solution générale y_h . Ceci nous donne trois degrés de liberté.

On dérive $y_p(x)$:

$$\begin{aligned} y_p'(x) &= [c_1'(x)y_1(x) + c_2'(x)y_2(x) + c_3'(x)y_3(x)] \\ &\quad + c_1(x)y_1'(x) + c_2(x)y_2'(x) + c_3(x)y_3'(x) \\ &= c_1(x)y_1'(x) + c_2(x)y_2'(x) + c_3(x)y_3'(x). \end{aligned}$$

où l'on a employé un 1er degré de liberté pour supposer que le terme entre crochets est nul :

$$c_1'(x)y_1(x) + c_2'(x)y_2(x) + c_3'(x)y_3(x) = 0. \quad (3.24)$$

On dérive $y_p'(x)$:

$$\begin{aligned} y_p''(x) &= [c_1'(x)y_1'(x) + c_2'(x)y_2'(x) + c_3'(x)y_3'(x)] \\ &\quad + c_1(x)y_1''(x) + c_2(x)y_2''(x) + c_3(x)y_3''(x) \\ &= c_1(x)y_1''(x) + c_2(x)y_2''(x) + c_3(x)y_3''(x), \end{aligned}$$

où l'on a employé un 2ème degré de liberté pour supposer que le terme entre crochets est nul :

$$c'_1(x)y'_1(x) + c'_2(x)y'_2(x) + c'_3(x)y'_3(x) = 0. \quad (3.25)$$

Enfin

$$y_p'''(x) = [c'_1(x)y_1''(x) + c'_2(x)y_2''(x) + c'_3(x)y_3''(x)] \\ + [c_1(x)y_1'''(x) + c_2(x)y_2'''(x) + c_3(x)y_3'''(x)].$$

Utilisant les expressions obtenues pour y_p , y_p' , y_p'' et y_p''' , on a

$$Ly_p = c'_1y_1'' + c'_2y_2'' + c'_3y_3'' \\ + [c_1Ly_1 + c_2Ly_2 + c_3Ly_3] \\ = c'_1y_1'' + c'_2y_2'' + c'_3y_3'' \\ = r(x),$$

puisque y_1 , y_2 et y_3 sont des solutions de $Ly = 0$ et donc le terme entre crochets est nul. De plus, on cherche y_p telle que $Ly_p = r(x)$; ceci utilise le 3ème degré de liberté. On a donc

$$c'_1y_1'' + c'_2y_2'' + c'_3y_3'' = r(x). \quad (3.26)$$

On réécrit les trois équations (3.24)–(3.26) en $c'_1(x)$, $c'_2(x)$ et $c'_3(x)$ sous forme matricielle

$$\begin{bmatrix} y_1(x) & y_2(x) & y_3(x) \\ y_1'(x) & y_2'(x) & y_3'(x) \\ y_1''(x) & y_2''(x) & y_3''(x) \end{bmatrix} \begin{bmatrix} c'_1(x) \\ c'_2(x) \\ c'_3(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ r(x) \end{bmatrix}, \quad (3.27)$$

c'est-à-dire

$$A(x)\mathbf{c}'(x) = \begin{bmatrix} 0 \\ 0 \\ r(x) \end{bmatrix}.$$

Puisque y_1 , y_2 et y_3 forment un système fondamental, il suit que

$$\det A = W(y_1, y_2, y_3) \neq 0,$$

le wronskien n'étant jamais nul d'après le corollaire 3.1.

On résout le système linéaire pour $\mathbf{c}'(x)$ et l'on intègre

$$\mathbf{c}(x) = \int \mathbf{c}'(x) dx.$$

Aucune constante d'intégration n'est requise ici puisque la solution générale en contiendra trois. La solution générale de (3.21) est donc

$$y_g(x) = Ay_1 + By_2 + Cy_3 + c_1(x)y_1 + c_2(x)y_2 + c_3(x)y_3. \quad (3.28)$$

La méthode de Cramer produit une jolie formule pour résoudre le système (3.27) en 2 et 3 dimensions à cause de la forme particulière du second membre. En 2 dimensions, on a :

$$y_p(x) = -y_1(x) \int \frac{y_2(x)r(x)}{W(x)} dx + y_2(x) \int \frac{y_1(x)r(x)}{W(x)} dx. \quad (3.29)$$

En 3 dimensions on a :

$$c'_1(x) = \frac{r}{W} \begin{vmatrix} y_2 & y_3 \\ y_2' & y_3' \end{vmatrix}, \quad c'_2(x) = -\frac{r}{W} \begin{vmatrix} y_1 & y_3 \\ y_1' & y_3' \end{vmatrix}, \quad c'_3(x) = \frac{r}{W} \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}. \quad (3.30)$$

Il ne reste qu'à intégrer les c'_i par rapport à x et à écrire $y_p(x)$ comme à l'équation (3.23).

REMARQUE 3.3. Si le coefficient $a_n(x)$ de $y^{(n)}$ n'est pas égal à 1, il faut diviser le 2ème membre de (3.27) par $a_n(x)$, c'est-à-dire remplacer $r(x)$ par $r(x)/a_n(x)$.

EXEMPLE 3.14. Trouver la solution générale de l'équation différentielle

$$(D^2 + 1)y = \sec x \tan x,$$

par variation des paramètres.

RÉSOLUTION. (a) **Résolution analytique.** — Puisque le 2ème membre n'admet pas un nombre fini de dérivées dépendantes, la méthode des coefficients indéterminés ne fonctionne pas.

On sait que la solution générale de l'équation homogène est

$$y_h(x) = c_1 \cos x + c_2 \sin x.$$

On cherche une solution particulière de l'équation non homogène par variation des paramètres :

$$y_p(x) = c_1(x) \cos x + c_2(x) \sin x.$$

On a donc

$$\begin{bmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{bmatrix} \begin{bmatrix} c'_1(x) \\ c'_2(x) \end{bmatrix} = \begin{bmatrix} 0 \\ \sec x \tan x \end{bmatrix},$$

c'est-à-dire

$$Q(x)\mathbf{c}'(x) = \begin{bmatrix} 0 \\ \sec x \tan x \end{bmatrix}. \quad (3.31)$$

Puisque la matrice Q est orthogonale, c'est-à-dire

$$QQ^T = \begin{bmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{bmatrix} \begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

il suit que l'inverse Q^{-1} de Q est la transposée Q^T de Q :

$$Q^{-1} = Q^T.$$

On obtient \mathbf{c}' en multipliant (3.31) à gauche par Q^T :

$$\mathbf{c}' = Q^T Q \mathbf{c}' = Q^T \begin{bmatrix} 0 \\ \sec x \tan x \end{bmatrix},$$

c'est-à-dire

$$\begin{bmatrix} c'_1(x) \\ c'_2(x) \end{bmatrix} = \begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix} \begin{bmatrix} 0 \\ \sec x \tan x \end{bmatrix}.$$

Alors

$$\begin{aligned} c'_1 &= -\frac{\sin x}{\cos x} \tan x = -\tan^2 x, \\ c'_2 &= \frac{\cos x}{\cos x} \tan x = \tan x = \frac{\sin x}{\cos x}, \end{aligned}$$

qu'on intègre :

$$\begin{aligned} c_1 &= -\int (\sec^2 x - 1) dx = x - \tan x, \\ c_2 &= -\ln |\cos x| = \ln |\sec x|. \end{aligned}$$

La solution particulière est donc

$$y_p(x) = (x - \tan x) \cos x + (\ln |\sec x|) \sin x$$

et la solution générale est

$$\begin{aligned} y(x) &= y_h(x) + y_p(x) \\ &= A \cos x + B \sin x + (x - \tan x) \cos x + (\ln |\sec x|) \sin x. \end{aligned}$$

(b) Résolution par Matlab symbolique.—

`dsolve('D2y+y=sec(x)*tan(x)', 'x')`

`y = -log(cos(x))*sin(x)-sin(x)+x*cos(x)+C1*sin(x)+C2*cos(x)`

□

EXEMPLE 3.15. Trouver la solution générale de l'équation différentielle

$$y''' - y' = \cosh x,$$

par variation des paramètres.

RÉSOLUTION. L'équation caractéristique :

$$\lambda^3 - \lambda = \lambda(\lambda^2 - 1) = 0 \implies \lambda_1 = 0, \lambda_2 = 1, \lambda_3 = -1.$$

La solution générale de l'équation homogène est

$$y_h(x) = c_1 + c_2 e^x + c_3 e^{-x}.$$

Par variation des paramètres, la solution particulière de l'équation non homogène est

$$y_p(x) = c_1(x) + c_2(x) e^x + c_3(x) e^{-x}.$$

On a donc le système

$$\begin{bmatrix} 1 & e^x & e^{-x} \\ 0 & e^x & -e^{-x} \\ 0 & e^x & e^{-x} \end{bmatrix} \begin{bmatrix} c_1'(x) \\ c_2'(x) \\ c_3'(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cosh x \end{bmatrix}.$$

On résout ce système par élimination gaussienne :

$$\begin{bmatrix} 1 & e^x & e^{-x} \\ 0 & e^x & -e^{-x} \\ 0 & 0 & 2e^{-x} \end{bmatrix} \begin{bmatrix} c_1'(x) \\ c_2'(x) \\ c_3'(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cosh x \end{bmatrix}.$$

Alors

$$c_3' = \frac{1}{2} e^x \cosh x = \frac{1}{2} e^x \left(\frac{e^x + e^{-x}}{2} \right) = \frac{1}{4} (e^{2x} + 1),$$

$$c_2' = e^{-2x} c_3' = \frac{1}{4} (1 + e^{-2x}),$$

$$c_1' = -e^x c_2' - e^{-x} c_3' = -\frac{1}{2} (e^x + e^{-x}) = -\cosh x.$$

On intègre :

$$c_1 = -\sinh x$$

$$c_2 = \frac{1}{4} \left(x - \frac{1}{2} e^{-2x} \right)$$

$$c_3 = \frac{1}{4} \left(\frac{1}{2} e^{2x} + x \right).$$

La solution particulière est

$$\begin{aligned} y_p(x) &= -\sinh x + \frac{1}{4} \left(x e^x - \frac{1}{2} e^{-x} \right) + \frac{1}{4} \left(\frac{1}{2} e^x + x e^{-x} \right) \\ &= -\sinh x + \frac{1}{4} x (e^x + e^{-x}) + \frac{1}{8} (e^x - e^{-x}) \\ &= \frac{1}{2} x \cosh x - \frac{3}{4} \sinh x. \end{aligned}$$

La solution générale de l'équation non homogène est donc

$$\begin{aligned} y_g(x) &= A + B' e^x + C' e^{-x} + \frac{1}{2} x \cosh x - \frac{3}{4} \sinh x \\ &= A + B e^x + C e^{-x} + \frac{1}{2} x \cosh x, \end{aligned}$$

puisque la fonction

$$\sinh x = \frac{e^x - e^{-x}}{2}$$

est déjà contenue dans la solution générale y_h de l'équation homogène. Matlab symbolique ne produit pas une solution générale aussi simple. \square

Si l'on emploie la méthode des coefficients indéterminés pour résoudre ce problème, il faut prendre une solution particulière de la forme

$$y_p(x) = ax \cosh x + bx \sinh x,$$

puisque $\cosh x$ et $\sinh x$ sont des combinaisons linéaires de e^x et e^{-x} qui sont des solutions de l'équation homogène. En effet, si l'on pose

$$y_p(x) = ax \cosh x + bx \sinh x$$

dans l'équation $y''' - y' = \cosh x$, on obtient

$$\begin{aligned} y_p''' - y_p' &= 2a \cosh x + 2b \sinh x \\ &= \cosh x, \end{aligned}$$

d'où

$$a = \frac{1}{2} \quad \text{et} \quad b = 0.$$

EXEMPLE 3.16. Trouver la solution générale de l'équation différentielle

$$Ly := y'' + 3y' + 2y = \frac{1}{1 + e^x}.$$

RÉSOLUTION. Puisque la dimension de l'espace des dérivées du 2e membre est infinie, on emploie la méthode de la variation des paramètres.

On remarque que la commande `dsolve` de Matlab symbolique produit une longue solution inutilisable. On suit donc la méthode théorique de Lagrange, mais on exécute les simples manipulations d'algèbre et de calcul infinitésimal au moyen de Matlab.

Le polynôme caractéristique de l'équation homogène $Ly = 0$ est

$$\lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2) = 0 \implies \lambda_1 = -1, \quad \lambda_2 = -2.$$

Donc, la solution générale $y_h(x)$ de $Ly = 0$ est

$$y_h(x) = c_1 e^{-x} + c_2 e^{-2x}.$$

Par variation des paramètres, on cherche une solution particulière de l'équation non homogène de la forme

$$y_p(x) = c_1(x) e^{-x} + c_2(x) e^{-2x}.$$

Les fonctions $c_1(x)$ et $c_2(x)$ sont les intégrales des solutions $c'_1(x)$ et $c'_2(x)$ du système algébrique $A\mathbf{c}' = \mathbf{b}$,

$$\begin{bmatrix} e^{-x} & e^{-2x} \\ -e^{-x} & -2e^{-2x} \end{bmatrix} \begin{bmatrix} c'_1 \\ c'_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/(1+e^x) \end{bmatrix}.$$

On utilise Matlab symbolique pour résoudre ce petit système.

```
>> clear
>> syms x real; syms c dc A b yp;
>> A = [exp(-x) exp(-2*x); -exp(-x) -2*exp(-2*x)];
>> b=[0 1/(1+exp(x))]' ;
>> dc = A\b % solve for c'(x)
dc =
 [ 1/exp(-x)/(1+exp(x))]
 [-1/exp(-2*x)/(1+exp(x))]
>> c = int(dc) % c(x) est l'integrale de c'(x)
c =
 [ log(1+exp(x))]
 [-exp(x)+log(1+exp(x))]
>> yp=c'*[exp(-x) exp(-2*x)]'
yp =
 log(1+exp(x))*exp(-x)+(-exp(x)+log(1+exp(x)))*exp(-2*x)
```

Puisque $-e^{-x}$ est contenu dans $y_h(x)$, la solution générale de l'équation non homogène est

$$y(x) = A e^{-x} + B e^{-2x} + [\ln(1+e^x)] e^{-x} + [\ln(1+e^x)] e^{-2x}. \quad \square$$

EXEMPLE 3.17. Résoudre le problème à valeur initiale de l'équation d'Euler-Cauchy non homogène :

$$Ly := 2x^2 y'' + xy' - 3y = x^{-3}, \quad y(1) = 0, \quad y'(1) = 2.$$

RÉSOLUTION. L'insertion de $y = x^m$ dans l'équation homogène $Ly = 0$ donne le polynôme caractéristique :

$$2m^2 - m - 3 = 0 \implies m_1 = \frac{3}{2}, \quad m_2 = -1.$$

Alors la solution générale $y_h(x)$ de $Ly = 0$ est

$$y_h(x) = c_1 x^{3/2} + c_2 x^{-1}.$$

On trouve une solution particulière $y_p(x)$ de l'équation non homogène par la méthode de la variation des paramètres puisque la dimension de l'espace des dérivées du second membre est infinie. Posons :

$$y_p(x) = c_1(x) x^{3/2} + c_2(x) x^{-1}.$$

Il nous faut résoudre le système linéaire :

$$\begin{bmatrix} x^{3/2} & x^{-1} \\ \frac{3}{2} x^{1/2} & -x^{-2} \end{bmatrix} \begin{bmatrix} c'_1 \\ c'_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} x^{-5} \end{bmatrix}.$$

On a divisé le second membre par le coefficient $2x^2$ de y'' pour que l'équation soit sous forme standard avec le coefficient de y'' égal à 1. On résout ce système pour c'_1 et c'_2 :

$$c'_1 = \frac{1}{5} x^{-11/2}, \quad c'_2 = -\frac{1}{5} x^{-3}.$$

Et après une intégration :

$$c_1(x) = -\frac{2}{45} x^{-9/2}, \quad c_2(x) = \frac{1}{10} x^{-2}.$$

Enfin, la solution générale est

$$\begin{aligned} y(x) &= Ax^{3/2} + Bx^{-1} - \frac{2}{45} x^{-3} + \frac{1}{10} x^{-3} \\ &= Ax^{3/2} + Bx^{-1} + \frac{1}{18} x^{-3}. \end{aligned}$$

On détermine les constantes A and B par les conditions initiales. Pour ce faire, il faut la dérivée $y'(x)$ de $y(x)$:

$$y'(x) = \frac{3}{2} Ax^{1/2} - Bx^{-2} - \frac{1}{6} x^{-4}.$$

Alors,

$$\begin{aligned} y(1) &= A + B + \frac{1}{18} = 0, \\ y'(1) &= \frac{3}{2} A - B - \frac{1}{6} = 2. \end{aligned}$$

On résout pour A et B :

$$A = \frac{38}{45}, \quad B = -\frac{9}{10}.$$

La solution (unique) est

$$y(x) = \frac{38}{45} x^{3/2} - \frac{9}{10} x^{-1} + \frac{1}{18} x^{-3}. \quad \square$$

3.6. Systèmes asservis

On présente deux exemples de systèmes mécaniques asservis.

Soit une ressort en position verticale pendante fixé à une poutre rigide. Le ressort résiste à l'extension et à la compression et sa constante de Hooke est k . On étudie le mouvement vertical amorti et asservi d'une masse de m kg fixée au bout inférieur du ressort (V. figure 3.2). La constante d'amortissement est c et la force d'asservissement est $r(t)$.

On se réfère à l'exemple 2.5 pour la dérivation de l'équation différentielle qui gouverne le système non asservi, auquel on ajoute au 2ème membre la fonction d'asservissement :

$$y'' + \frac{c}{m} y' + \frac{k}{m} y = \frac{1}{m} r(t).$$

EXEMPLE 3.18 (Oscillation asservie sans résonance). Trouver la solution du problème asservi

$$Ly := y'' + 9y = 8 \sin t, \quad y(0) = 1, \quad y'(0) = 1,$$

et tracer la solution sur $[0, 7]$.

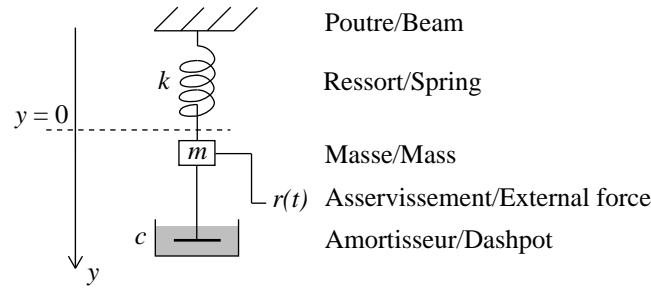


FIGURE 3.2. Système amorti et asservi.

RÉSOLUTION. **(a) Résolution analytique.**— La solution générale de $Ly = 0$ est

$$y_h(t) = A \cos 3t + B \sin 3t.$$

Par la méthode des coefficients indéterminés, on prend

$$y_p(t) = a \cos t + b \sin t,$$

qu'on substitue dans $Ly = 8 \sin t$. On obtient

$$\begin{aligned} y_p'' + 9y_p &= (-a + 9a) \cos t + (-b + 9b) \sin t \\ &= 8 \sin t. \end{aligned}$$

En identifiant les coefficients des deux derniers membres, on a

$$a = 0, \quad b = 1.$$

La solution générale de $Ly = 8 \sin t$ est

$$y(t) = A \cos 3t + B \sin 3t + \sin t.$$

On détermine A et B au moyen des conditions initiales :

$$\begin{aligned} y(0) &= A = 1, \\ y'(t) &= -3A \sin 3t + 3B \cos 3t + \cos t, \\ y'(0) &= 3B + 1 = 1 \implies B = 0. \end{aligned}$$

La solution, qui est unique, est donc

$$y(t) = \cos 3t + \sin t.$$

(b) Résolution par Matlab symbolique.—

```
dsolve('D2y+9*y=8*sin(t)', 'y(0)=1', 'Dy(0)=1', 't')
y = sin(t)+cos(3*t)
```

(c) Résolution par Matlab numérique.— Pour récrire l'équation différentielle du 2e ordre en système du 1er ordre on utilise les variables

$$\begin{aligned} y_1 &= y, \\ y_2 &= y', \end{aligned}$$

Alors,

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= -9y_1 + 8 \sin t. \end{aligned}$$

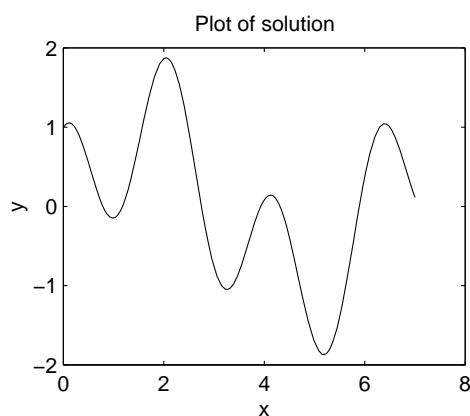


FIGURE 3.3. Graphe de la solution de l'équation linéaire de l'exemple 3.18.

Le fichier M `exp312.m` :

```
function yprime = exp312(t,y);
yprime = [y(2); -9*y(1)+8*sin(t)];
```

On appelle le solveur `ode23` et la commande `plot` :

```
tspan = [0 7]; % solution sur 0<=t<=7
y0 = [1; 1]; % conditions initiales
[x,y] = ode23('exp312',tspan,y0);
plot(x,y(:,1))
```

Le graphe de la solution numérique est dans la fig. 3.3. □

EXEMPLE 3.19 (Oscillation asservie avec résonance). Trouver la solution du problème asservi :

$$Ly := y'' + 9y = 6 \sin 3t, \quad y(0) = 1, \quad y'(0) = 2,$$

et tracer la solution sur $[0, 7]$.

RÉSOLUTION. (a) **Résolution analytique.**— La solution générale de $Ly = 0$ est

$$y_h(t) = A \cos 3t + B \sin 3t.$$

Puisque le 2^{ème} membre de $Ly = 6 \sin 3t$ est contenu dans y_h , pour la méthode des coefficients indéterminés, on prend

$$y_p(t) = at \cos 3t + bt \sin 3t.$$

On obtient alors

$$\begin{aligned} y_p'' + 9y_p &= -6a \sin 3t + 6b \cos 3t \\ &= 6 \sin 3t. \end{aligned}$$

En identifiant les coefficients des deux derniers membres, on a

$$a = -1, \quad b = 0.$$

La solution générale de $Ly = 6 \sin 3t$ est

$$y(t) = A \cos 3t + B \sin 3t - t \cos 3t.$$

On détermine A et B au moyen des conditions initiales :

$$y(0) = A = 1,$$

$$y'(t) = -3A \sin 3t + 3B \cos 3t - \cos 3t + 3t \sin 3t,$$

$$y'(0) = 3B - 1 = 2 \implies B = 1.$$

La solution, qui est unique, est donc

$$y(t) = \cos 3t + \sin 3t - t \cos 3t.$$

Le terme $-t \cos 3t$, dont l'amplitude croît, provient de la résonance du système parce que la fréquence de la force d'asservissement et celle du système coïncident.

(b) Résolution par Matlab symbolique.—

```
dsolve('D2y+9*y=6*sin(3*t)', 'y(0)=1', 'Dy(0)=2', 't')
y = sin(3*t)-cos(3*t)*t+cos(3*t)
```

(c) Résolution par Matlab numérique.— Pour récrire l'équation différentielle du 2e ordre en système du 1er ordre on utilise les variables

$$y_1 = y,$$

$$y_2 = y',$$

Alors,

$$y_1' = y_2,$$

$$y_2' = -9y_1 + 6 \sin 3t.$$

Le fichier M `exp313.m` :

```
function yprime = exp313(t,y);
yprime = [y(2); -9*y(1)+6*sin(3*t)];
```

On appelle le solveur `ode23` et la commande `plot` :

```
tspan = [0 7]; % solution sur 0<=t<=7
y0 = [1; 1]; % conditions initiales
[x,y] = ode23('exp313',tspan,y0);
plot(x,y(:,1))
```

Le graphe de la solution numérique est dans la fig. 3.4. □

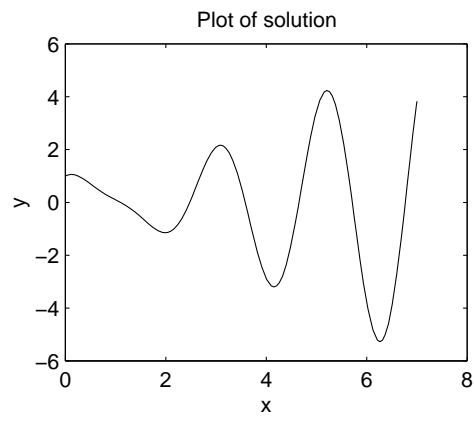


FIGURE 3.4. Graphe de la solution de l'équation linéaire de l'exemple 3.19.

Systèmes d'équations différentielles linéaires

4.1. Introduction

À la section 3.1, on a vu comment transformer une équation différentielle linéaire d'ordre n ,

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = r(x),$$

en un système linéaire de n équations du 1er ordre de la forme

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{bmatrix}' = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ r(x) \end{bmatrix},$$

avec n variables dépendantes :

$$u_1 = y, \quad u_2 = y', \quad \dots, \quad u_n = y^{(n-1)}.$$

Dans ce cas, les n valeurs initiales :

$$y(x_0) = k_1, \quad y'(x_0) = k_2, \quad \dots, \quad y^{(n-1)}(x_0) = k_n,$$

et le second membre $r(x)$ deviennent

$$\begin{bmatrix} u_1(x_0) \\ u_2(x_0) \\ \vdots \\ u_{n-1}(x_0) \\ u_n(x_0) \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_{n-1} \\ k_n \end{bmatrix}, \quad \mathbf{g}(x) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ r(x) \end{bmatrix}.$$

On écrit le système sous forme matricielle :

$$\mathbf{u}'(x) = A(x)\mathbf{u}(x) + \mathbf{g}(x), \quad \mathbf{u}(x_0) = \mathbf{k}.$$

La matrice $A(x)$ est une *matrice companion*.

EXEMPLE 4.1. Écrire l'équation différentielle du second ordre :

$$y'' + 5y' - y = e^x \tag{4.1}$$

en système de deux équations.

SOLUTION. Posons

$$u_1 = y, \quad u_2 = y' = u_1'.$$

Alors $y'' = u_2'$ et l'équation (4.1) devient

$$u_2' + 5u_2 - u_1 = e^x,$$

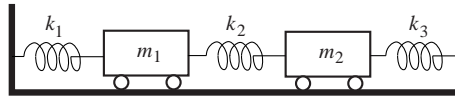


FIGURE 4.1. Système mécanique pour l'exemple 4.2.

c'est-à-dire

$$u_2' = u_1 - 5u_2 + e^x.$$

Si

$$\mathbf{u}(x) = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad \text{alors} \quad \mathbf{u}'(x) = \begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} u_2 \\ u_1 - 5u_2 + e^x \end{bmatrix}$$

ou bien

$$\mathbf{u}'(x) = \begin{bmatrix} 0 & 1 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 \\ e^x \end{bmatrix} = A\mathbf{u}(x) + \mathbf{g}(x). \quad \square$$

Dans ce chapitre, on considère des systèmes linéaires à n équations où la matrice $A(x)$ est une matrice $n \times n$ générale pas nécessairement une matrice compagnon, comme à l'exemple suivant.

EXEMPLE 4.2. Construire le système d'équations différentielles qui modélise le système mécanique de la fig. 4.1

RÉSOLUTION. Soit un système mécanique de deux masses m_1 et m_2 connectées entre elles par trois ressorts (V. fig. 4.1) de constantes de restauration (constante de Hooke) k_1 , k_2 et k_3 . On note par $x_1(t)$ et $x_2(t)$ la position des masses m_1 et m_2 à partir de leur point d'équilibre où $x > 0$ pointe vers la droite. Alors, $x_1''(t)$ et $x_2''(t)$ mesure l'accélération de chacune des masses. La résultante sur chaque masse provient de l'action des ressorts attachés à la masse. Chaque force est proportionnelle à la distance d'extension ou de compression du ressort. Par exemple, si la masse m_1 est à distance x_1 à droite de la position d'équilibre, l'action du ressort à gauche de m_1 est une force de restauration $-k_1x_1$ sur la masse pour ramener celle-ci à la position d'équilibre. L'action du ressort à droite de m_1 est une force de restauration $-k_2(x_2 - x_1)$ sur la masse pour ramener celle-ci à la position d'équilibre. Le terme k_2x_1 correspond à la compression du ressort du centre causée par la motion de m_1 , alors que le terme $-k_2x_2$ correspond à l'influence de la motion de m_2 sur ce ressort. Selon la seconde loi du mouvement de Newton, on obtient le système de deux équations couplées :

$$m_1x_1'' = -k_1x_1 + k_2(x_2 - x_1), \quad m_2x_2'' = -k_2(x_2 - x_1) - k_3x_2. \quad (4.2)$$

On transforme (4.2) en un système du 1er ordre au moyen des variables y_1 et y_2 qui représentent la vitesse des masses :

$$y_1 = x_1', \quad y_2 = x_2'. \quad (4.3)$$

Alors (4.2) devient un système de quatre équations à quatre inconnues x_1, y_1, x_2 et y_2 :

$$\begin{aligned} x_1' &= y_1, \\ y_1' &= \frac{-k_1 x_1 + k_2(x_2 - x_1)}{m_1}, \\ x_2' &= y_2, \\ y_2' &= \frac{-k_2(x_2 - x_1) - k_3 x_2}{m_2}, \end{aligned} \quad (4.4)$$

et sous forme matricielle :

$$\begin{bmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1+k_2}{m_1} & 0 & \frac{k_2}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_2}{m_2} & 0 & -\frac{k_2+k_3}{m_2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{bmatrix}. \quad (4.5)$$

Notons le vecteur inconnu, la matrice du système et les conditions initiales sous forme vectorielle :

$$\mathbf{u} = \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1+k_2}{m_1} & 0 & \frac{k_2}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_2}{m_2} & 0 & -\frac{k_2+k_3}{m_2} & 0 \end{bmatrix} \quad \mathbf{u}_0 = \begin{bmatrix} x_1(0) \\ y_1(0) \\ x_2(0) \\ y_2(0) \end{bmatrix}.$$

Alors, le problème aux valeurs initiales devient :

$$\mathbf{u}' = A\mathbf{u}, \quad \mathbf{u}(0) = \mathbf{u}_0. \quad (4.6)$$

On voit que A n'est pas une matrice compagnon. \square

4.2. Théorème d'existence et d'unicité

Dans cette section, on rappelle des résultats sur les systèmes cités dans les chapitres antérieurs. En particulier, le théorème 1.3 d'existence et d'unicité demeure valide pour les systèmes du 1er ordre de la forme

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y}), \quad \mathbf{y}(x_0) = \mathbf{y}_0, \quad (4.7)$$

pourvu que dans la définition 1.3 on emploie la norme au lieu du module dans la condition de Lipschitz :

$$\|\mathbf{f}(\mathbf{z}) - \mathbf{f}(\mathbf{y})\| \leq M\|\mathbf{z} - \mathbf{y}\|, \quad \text{pour tout } \mathbf{y}, \mathbf{z} \in \mathbb{R}^n,$$

et dans l'énoncé du théorème.

La même remarque s'applique au théorème 3.2 d'existence et d'unicité pour les systèmes linéaires de la forme

$$\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x), \quad \mathbf{y}(x_0) = \mathbf{y}_0, \quad (4.8)$$

pourvu que la matrice $A(x)$ et la fonction vectorielle $\mathbf{f}(x)$ soient continues sur l'intervalle (x_0, x_f) . La récurrence de Picard utilisée dans la démonstration de ce théorème est déjà énoncée pour les systèmes.

4.3. Système fondamental

On voit que les solutions d'un système linéaire homogène

$$\mathbf{y}' = A(x)\mathbf{y}, \quad x \in]a, b[, \quad (4.9)$$

forment un espace vectoriel puisque la différentiation et la multiplication sont des opérateurs linéaires.

Comme avant, m fonctions vectorielles $\mathbf{y}_1(x), \mathbf{y}_2(x), \dots, \mathbf{y}_m(x)$, sont *linéairement indépendantes* sur $]a, b[$ si l'identité

$$c_1\mathbf{y}_1(x) + c_2\mathbf{y}_2(x) + \dots + \mathbf{y}_m(x) = \mathbf{0}, \quad \text{pour tout } x \in]a, b[,$$

implique que

$$c_1 = c_2 = \dots = c_m = 0.$$

Sinon, ces fonctions sont *linéairement dépendantes*.

Pour un système général, le déterminant $W(x)$ de n fonctions vectorielles colonnes, $\mathbf{y}_1(x), \mathbf{y}_2(x), \dots, \mathbf{y}_n(x)$, à valeurs dans \mathbb{R}^n ,

$$W(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)(x) = \det \begin{bmatrix} y_{11}(x) & y_{12}(x) & \dots & y_{1n}(x) \\ y_{21}(x) & y_{22}(x) & \dots & y_{2n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1}(x) & y_{n2}(x) & \dots & y_{nn}(x) \end{bmatrix},$$

est une généralisation du wronskien pour une équation scalaire linéaire.

Pour énoncer et démontrer le lemme 3.1 de Liouville ou d'Abel pour un système linéaire général on définit la *trace* d'une matrice A , notée $\text{tr } A$, comme la somme des éléments a_{ii} de la diagonale de A ,

$$\text{tr } A = a_{11} + a_{22} + \dots + a_{nn}.$$

LEMME 4.1 (Abel). Soit $\mathbf{y}_1(x), \mathbf{y}_2(x), \dots, \mathbf{y}_n(x)$, n **solutions** du système $\mathbf{y}' = A(x)\mathbf{y}$ sur l'intervalle $]a, b[$. Alors, le déterminant $W(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)(x)$ satisfait l'identité suivante :

$$W(x) = W(x_0) e^{\int_{x_0}^x \text{tr } A(t) dt}, \quad x_0 \in]a, b[. \quad (4.10)$$

DÉMONSTRATION. Pour simplifier l'écriture on prend $n = 3$; le cas général se fait de la même manière. Soit $W(x)$ le déterminant de trois solutions $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$. Alors, la dérivée $W'(x)$ est de la forme

$$\begin{aligned} W'(x) &= \begin{vmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{vmatrix}' \\ &= \begin{vmatrix} y'_{11} & y'_{12} & y'_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{vmatrix} + \begin{vmatrix} y_{11} & y_{12} & y_{13} \\ y'_{21} & y'_{22} & y'_{23} \\ y_{31} & y_{32} & y_{33} \end{vmatrix} + \begin{vmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y'_{31} & y'_{32} & y'_{33} \end{vmatrix}. \end{aligned}$$

On considère le 1er déterminant du 3ème membre. On voit que la 1ère ligne du système différentiel

$$\begin{bmatrix} y'_{11} & y'_{12} & y'_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix}$$

est

$$\begin{aligned}y'_{11} &= a_{11}y_{11} + a_{12}y_{21} + a_{13}y_{31}, \\y'_{12} &= a_{11}y_{12} + a_{12}y_{22} + a_{13}y_{32}, \\y'_{13} &= a_{11}y_{13} + a_{12}y_{23} + a_{13}y_{33}.\end{aligned}$$

Si l'on substitue ces expressions dans la 1ère ligne du 1er déterminant et si de la 1ère ligne on soustrait a_{12} fois la 2ème ligne et a_{13} fois la 3ème ligne, on obtient $a_{11}W(x)$. De même pour le 2ème et le 3ème déterminants on obtient respectivement $a_{22}W(x)$ et $a_{33}W(x)$. Donc $W(x)$ satisfait l'équation séparable

$$W'(x) = \operatorname{tr}(A(x))W(x)$$

qui admet la solution

$$W(x) = W(x_0) e^{\int_{x_0}^x \operatorname{tr} A(t) dt}. \quad \square$$

Le corollaire suivant découle du lemme d'Abel.

COROLLAIRE 4.1. *Si n solutions du système différentiel homogène (4.9) sont indépendantes en un point, alors elles sont indépendantes sur l'intervalle $]a, b[$. Si, d'autre part, ces solutions sont linéairement dépendantes en un point, alors leur déterminant, $W(x)$, est identiquement nul, par conséquent elles sont dépendantes partout.*

REMARQUE 4.1. Il y a une différence entre l'indépendance linéaire de fonctions vectorielles et de solutions de systèmes linéaires. Par exemple, les deux fonctions vectorielles

$$\mathbf{f}_1(x) = \begin{bmatrix} x \\ 0 \end{bmatrix}, \quad \mathbf{f}_2(x) = \begin{bmatrix} 1+x \\ 0 \end{bmatrix},$$

sont linéairement indépendantes, mais leur déterminant est nul. Ceci ne contredit pas le corollaire 4.1 parce que \mathbf{f}_1 et \mathbf{f}_2 ne peuvent être solutions d'un système (4.9).

DÉFINITION 4.1. Un ensemble de n solutions linéairement indépendantes d'un système linéaire homogène $\mathbf{y}' = A(x)\mathbf{y}$ est un *système fondamental* et la matrice inversible correspondante

$$Y(x) = \begin{bmatrix} y_{11}(x) & y_{12}(x) & \cdots & y_{1n}(x) \\ y_{21}(x) & y_{22}(x) & \cdots & y_{2n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1}(x) & y_{n2}(x) & \cdots & y_{nn}(x) \end{bmatrix},$$

est une *matrice fondamentale*.

LEMME 4.2. *Si $Y(x)$ est une matrice fondamentale, alors $Z(x) = Y(x)Y^{-1}(x_0)$ est aussi une matrice fondamentale telle que $Z(x_0) = I$.*

DÉMONSTRATION. Soit C une matrice quelconque. Puisque $Y' = AY$, il suit que $(YC)' = Y'C = (AY)C = A(YC)$. Le lemme suit en prenant $C = Y^{-1}(x_0)$. On voit que $Z(x_0) = I$. \square

Dans la suite on supposera souvent qu'une matrice fondamentale satisfait la condition $Y(x_0) = I$. On a le théorème suivant pour les systèmes linéaires homogènes.

THÉORÈME 4.1. Soit $Y(x)$ une matrice fondamentale du système $\mathbf{y}' = A(x)\mathbf{y}$. Alors la solution générale est

$$\mathbf{y}(x) = Y(x)\mathbf{c},$$

où le vecteur \mathbf{c} is arbitraire. Si $Y(x_0) = I$, alors

$$\mathbf{y}(x) = Y(x)\mathbf{y}_0$$

est la solution unique du problème à valeurs initiales

$$\mathbf{y}' = A(x)\mathbf{y}, \quad \mathbf{y}(x_0) = \mathbf{y}_0.$$

DÉMONSTRATION. La démonstration des deux énoncés fait appel au théorème d'unicité. Pour la 1ère partie, soit $Y(x)$ une matrice fondamentale et $\mathbf{z}(x)$ une solution quelconque du système. Soit x_0 dans le domaine de $\mathbf{z}(x)$ et \mathbf{c} défini par

$$\mathbf{c} = Y^{-1}(x_0)\mathbf{z}(x_0).$$

Prenons $\mathbf{y}(x) = Y(x)\mathbf{c}$. Puisque $\mathbf{y}(x)$ et $\mathbf{z}(x)$ satisfont la même équation différentielle avec les mêmes conditions initiales, ils sont deux solutions identiques par le théorème d'unicité. La démonstration de la 2ème partie est semblable. \square

On emploiera le lemme qui suit pour exprimer la solution du problème à valeurs initiales (4.8) au moyen d'une solution fondamentale.

LEMME 4.3. Soit $Y(x)$ une matrice fondamentale du système (4.9). Alors, $(Y^T)^{-1}(x)$ est une solution fondamentale du système adjoint

$$\mathbf{y}' = -A^T(x)\mathbf{y}. \quad (4.11)$$

DÉMONSTRATION. La dérivée de l'identité

$$Y^{-1}(x)Y(x) = I,$$

est

$$(Y^{-1})'(x)Y(x) + Y^{-1}(x)Y'(x) = 0.$$

Puisque la matrice $Y(x)$ est une solution de (4.9), on peut remplacer $Y'(x)$ du 2ème terme du 1er membre par $A(x)Y(x)$, ce qui donne

$$(Y^{-1})'(x)Y(x) = -Y^{-1}(x)A(x)Y(x).$$

En multipliant cette équation à droite par $Y^{-1}(x)$ et en prenant la transposée des deux membres on obtient (4.11). \square

THÉORÈME 4.2 (Formule de solution). Soit $Y(x)$ une matrice fondamentale solution du système linéaire homogène (4.9). Alors, la solution unique du problème aux valeurs initiales (4.8) est

$$\mathbf{y}(x) = Y(x)Y^{-1}(x_0)\mathbf{y}_0 + Y(x) \int_{x_0}^x Y^{-1}(t)\mathbf{g}(t) dt. \quad (4.12)$$

DÉMONSTRATION. On multiplie les deux membres de (4.8) par $Y^{-1}(x)$. Alors par le lemme 4.3 on obtient

$$(Y^{-1}(x)\mathbf{y}(x))' = Y^{-1}(x)\mathbf{g}(x).$$

On termine la démonstration en intégrant cette expression par rapport à x de x_0 à x . \square

4.4. Systèmes linéaires à coefficients constants

La solution d'un système linéaire à coefficients constants peut s'exprimer au moyen des valeurs propres et vecteurs propres de la matrice des coefficients A . Soit un système linéaire homogène :

$$\mathbf{y}' = A\mathbf{y}, \quad (4.13)$$

à matrice constante A d'ordre n . On cherche une solution de la forme

$$\mathbf{y}(x) = e^{\lambda x}\mathbf{v}, \quad (4.14)$$

où le nombre λ et le vecteur \mathbf{v} sont à déterminer. Si l'on substitue (4.14) dans (4.13) et l'on divise par $e^{\lambda x}$ on obtient le problème de valeur propre :

$$(A - \lambda I)\mathbf{v} = \mathbf{0}, \quad \mathbf{v} \neq \mathbf{0}. \quad (4.15)$$

Cette équation admet une solution non-nulle \mathbf{v} si et seulement si

$$\det(A - \lambda I) = 0.$$

Le 1er membre est un polynôme de degré n . Les n racines de cette équation sont les *valeurs propres* de A et les vecteurs correspondants non-nuls \mathbf{v} sont les *vecteurs propres* de A .

Pour chaque valeur propre distincte, A admet un vecteur propre et l'ensemble de ces vecteurs propres sont linéairement indépendants. Si A est symétrique, $A^T = A$, c'est-à-dire A et sa transposée sont égales, alors les valeurs propres sont réelles et A admet n vecteurs propres qu'on peut choisir orthonormés.

EXEMPLE 4.3. Trouver la solution générale du système symétrique $\mathbf{y}' = A\mathbf{y}$:

$$\mathbf{y}' = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \mathbf{y}.$$

RÉSOLUTION. On obtient les valeurs propres au moyen du polynôme caractéristique de A ,

$$\det(A - \lambda I) = \det \begin{bmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{bmatrix} = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3) = 0.$$

Donc

$$\lambda_1 = 1, \quad \lambda_2 = 3.$$

Le vecteur propre associé à λ_1 est solution du système singulier

$$(A - I)\mathbf{u} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \mathbf{0}.$$

Si l'on prend $u_1 = 1$, alors le vecteur propre est

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

De même, le vecteur propre associé à λ_2 est solution du système singulier

$$(A - 3I)\mathbf{v} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mathbf{0}.$$

Si l'on prend $v_1 = 1$, alors le vecteur propre est

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Puisque $\lambda_1 \neq \lambda_2$, on a deux solutions indépendantes :

$$\mathbf{y}_1 = e^x \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{y}_2 = e^{3x} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Le système fondamental et la solution générale sont

$$Y(x) = \begin{bmatrix} e^x & e^{3x} \\ -e^x & e^{3x} \end{bmatrix}, \quad \mathbf{y} = Y(x)\mathbf{c}.$$

La solution par Matlab est

```
A = [2 1; 1 2];
[Y,D] = eig(A);
syms x c1 c2
z = Y*diag(exp(diag(D*x)))*[c1; c2]
z =
[ 1/2*2^(1/2)*exp(x)*c1+1/2*2^(1/2)*exp(3*x)*c2]
[ -1/2*2^(1/2)*exp(x)*c1+1/2*2^(1/2)*exp(3*x)*c2]
```

On remarque que Matlab normalise les vecteurs propres dans la norme l_2 . Donc la matrice Y est orthogonale puisque A est symétrique. La solution \mathbf{y}

```
y = simplify(sqrt(2)*z)
y =
[ exp(x)*c1+exp(3*x)*c2]
[ -exp(x)*c1+exp(3*x)*c2]
```

est obtenue au moyen des vecteurs propres \mathbf{u} et \mathbf{v} non normalisés. \square

Si la matrice constante A du système $\mathbf{y}' = A\mathbf{y}$ admet un ensemble complet de vecteurs propres indépendants, alors elle est diagonalisable :

$$Y^{-1}AY = D.$$

Les colonnes de la matrice Y sont les vecteurs propres de A et les valeurs propres associées sont les éléments de la matrice diagonale D . On peut employer ce processus pour résoudre le problème aux valeurs initiales :

$$\mathbf{y}' = A\mathbf{y}, \quad \mathbf{y}(0) = \mathbf{y}_0.$$

Posons

$$\mathbf{y} = Y\mathbf{x}, \quad \text{ou} \quad \mathbf{x} = Y^{-1}\mathbf{y},$$

Puisque la matrice A est constante, alors Y est constante et $\mathbf{x}' = Y^{-1}\mathbf{y}'$. Ceci transforme le système $\mathbf{y}' = A\mathbf{y}$:

$$Y^{-1}\mathbf{y}' = Y^{-1}AYY^{-1}\mathbf{y},$$

c'est-à-dire

$$\mathbf{x}' = D\mathbf{x}.$$

Les équations se découpent :

$$x_1'(t) = \lambda_1 x_1(t), \quad x_2'(t) = \lambda_2 x_2(t), \quad \dots, \quad x_n(t)' = \lambda_n x_n(t),$$

et l'on obtient la solution par une intégration :

$$x_1(t) = c_1 e^{\lambda_1 t}, \quad x_2(t) = c_2 e^{\lambda_2 t}, \quad \dots, \quad x_n(t) = c_n e^{\lambda_n t}.$$

Les conditions initiales déterminent les constantes c_1, \dots, c_n . Puisque

$$\mathbf{y}_0 = Y\mathbf{x}(0) = Y\mathbf{c},$$

alors

$$\mathbf{c} = Y^{-1}\mathbf{y}_0.$$

On utilise ces résultats à l'exemple suivant.

EXEMPLE 4.4. Résoudre le système de l'exemple 4.2 avec

$$m_1 = 10, \quad m_2 = 20, \quad k_1 = k_2 = k_3 = 1,$$

et les valeurs initiales

$$x_1 = 0.8, \quad x_2 = y_0 = y_1 = 0.$$

Tracer la solution.

RÉSOLUTION. La matrice A est de la forme

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.2 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 1 \\ 0.05 & 0 & -0.1 & 0 \end{bmatrix}$$

La programme de Matlab trouve la solution $x(t)$ et $y(t)$ et le graphique :

```
A = [0 1 0 0; -0.2 0 0.1 0; 0 0 0 1; 0.05 0 -0.1 0];
y0 = [0.8 0 0 0]';
[Y,D] = eig(A);
t = 0:1/5:60; c = inv(Y)*y0; y = y0;
for i = 1:length(t)-1
yy = Y*diag(exp(diag(D)*t(i+1)))*c;
y = [y,yy];
end
ry = real(y); % la solution est réelle;
subplot(2,2,1); plot(t,ry(1,:),t,ry(3,:), '--');
```

La commande `ode45` de la suite `ode` de Matlab produit la solution numérique. On emploie le fichier M `spring.m`,

```
function yprime = spring(t,y); % MAT 2784, Exemple 3a.4.2.
A = [0 1 0 0; -0.2 0 0.1 0; 0 0 0 1; 0.05 0 -0.1 0];
yprime = A*y;
```

Alors

```
y0 = [0.8 0 0 0]; tspan=[0 60];
[t,y]=ode45('spring',tspan,y0);
subplot(2,2,1); plot(t,y(:,1),t,y(:,3));
```

Le graphique se trouve à la fig. 4.2. □

Si, dans le cas de valeurs propres multiples, on manque de vecteurs propres on recourt aux vecteurs propres généralisés.

DÉFINITION 4.2. Soit A une matrice d'ordre n . On dit qu'une valeur propre λ de A de multiplicité $m > 1$ est *déficiente* si A admet $k < m$ vecteurs propres indépendants associés à λ . Dans ce cas le nombre

$$r = m - k$$

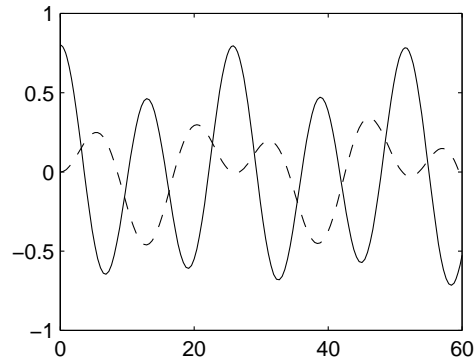


FIGURE 4.2. Graphique de la solution $x_1(t)$ (ligne pleine) et $x_2(t)$ (ligne hachurée) de l'exemple 4.4.

est le *degré de déficience* de λ . Un vecteur \mathbf{u} est un *vecteur propre généralisé* de A associé à λ s'il existe un entier $s > 1$ tel que

$$(A - \lambda I)^s \mathbf{u} = \mathbf{0},$$

mais

$$(A - \lambda I)^{s-1} \mathbf{u} \neq \mathbf{0}.$$

Soit une matrice A qui admet une valeur propre λ déficiente de degré r et un vecteur propre associé \mathbf{u}_1 . On construit les vecteurs propres généralisés $\{\mathbf{u}_2, \dots, \mathbf{u}_r\}$ solutions des systèmes

$$(A - \lambda I)\mathbf{u}_2 = \mathbf{u}_1, \quad (A - \lambda I)\mathbf{u}_3 = \mathbf{u}_2, \quad \dots, \quad (A - \lambda I)\mathbf{u}_r = \mathbf{u}_{r-1}.$$

Le vecteur propre \mathbf{u}_1 et les vecteurs propres généralisés engendrent les solutions linéairement indépendantes de (4.13) :

$$\mathbf{y}_1(x) = e^{\lambda x} \mathbf{u}_1, \quad \mathbf{y}_2(x) = e^{\lambda x} (x\mathbf{u}_1 + \mathbf{u}_2), \quad \mathbf{y}_3(x) = e^{\lambda x} \left(\frac{x^2}{2} \mathbf{u}_1 + x\mathbf{u}_2 + \mathbf{u}_3 \right), \quad \dots$$

On démontre en algèbre linéaire qu'une matrice quelconque d'ordre n admet n vecteurs propres généralisés indépendants.

Regardons en détail la situation pour un système 2×2 . Supposons que $\mathbf{y}' = A\mathbf{y}$ admet une valeur propre double, $\lambda_1 = \lambda_2 = \lambda$. Alors une solution est

$$\mathbf{y}_1(x) = e^{\lambda x} \mathbf{v}.$$

D'après le chapitre 2, pourrait croire que la seconde solution serait de la forme

$$\mathbf{y}_2(x) = x e^{\lambda x} \mathbf{v}.$$

Dans ce cas on a

$$\mathbf{y}'_2(x) = e^{\lambda x} \mathbf{v} + \lambda x e^{\lambda x} \mathbf{v},$$

alors que

$$A\mathbf{y}_2(x) = Ax e^{\lambda x} \mathbf{v} = x e^{\lambda x} A\mathbf{v} = x e^{\lambda x} \mathbf{v}.$$

Il suit que $\mathbf{y}_2(x)$ n'est pas une solution ($\mathbf{y}'_2 \neq A\mathbf{y}_2$). Cherchons une solution de la forme

$$\mathbf{y}_2(x) = e^{\lambda x} (x\mathbf{v} + \mathbf{u}).$$

Alors,

$$\mathbf{y}'_2(x) = e^{\lambda x} \mathbf{v} + \lambda x e^{\lambda x} \mathbf{v} + \lambda e^{\lambda x} \mathbf{u}$$

et

$$A\mathbf{y}_2 = A(x e^{\lambda x} \mathbf{v} + e^{\lambda x} \mathbf{u}) = x\lambda e^{\lambda x} \mathbf{v} + e^{\lambda x} A\mathbf{u}.$$

Donc, après simplifications, on voit que $\mathbf{y}'_2 = A\mathbf{y}_2$ exige que

$$e^{\lambda x} \mathbf{v} + \lambda x e^{\lambda x} \mathbf{v} + \lambda e^{\lambda x} \mathbf{u} = x\lambda e^{\lambda x} \mathbf{v} + e^{\lambda x} A\mathbf{u}$$

$$e^{\lambda x} \mathbf{v} + \lambda e^{\lambda x} \mathbf{u} = e^{\lambda x} A\mathbf{u}$$

$$\mathbf{v} + \lambda \mathbf{u} = A\mathbf{u},$$

c'est-à-dire

$$(A - \lambda I)\mathbf{u} = \mathbf{v}. \quad (4.16)$$

Si \mathbf{u} satisfait (4.16), alors

$$\mathbf{y}_2(x) = e^{\lambda x} (x\mathbf{v} + \mathbf{u})$$

est une seconde solution indépendante.

EXEMPLE 4.5. Trouver la solution générale du système

$$\mathbf{y}' = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{y}.$$

RÉSOLUTION. Puisque

$$\det(A - \lambda I) = \left| \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \left| \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} \right| = (1-\lambda)^2 = 0,$$

on a une valeur propre répétée $\lambda_1 = \lambda_2 = 1$. Le vecteur propre \mathbf{v} doit satisfaire le système homogène $(A - \lambda I)\mathbf{v} = \mathbf{0}$:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{v} = \mathbf{0}$$

Si l'on prend

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

la première solution est

$$\mathbf{y}_1(x) = e^{\lambda x} \mathbf{v} = e^x \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Pour obtenir la seconde solution, on résout le système nonhomogène

$$(A - \lambda I)\mathbf{u} = \mathbf{v}$$

pour \mathbf{u} , c'est-à-dire

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Si l'on prend

$$\mathbf{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

la seconde solution est

$$\mathbf{y}_2(x) = e^{\lambda x} (x\mathbf{v} + \mathbf{u}) = e^x \left(x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right).$$

La solution générale est donc

$$\mathbf{y}_g(x) = c_1 \mathbf{y}_1(x) + c_2 \mathbf{y}_2(x) = c_1 e^x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^x \left(x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right). \quad \square$$

EXEMPLE 4.6. Résoudre le système $\mathbf{y}' = A\mathbf{y}$:

$$\mathbf{y}' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix} \mathbf{y}.$$

RÉSOLUTION. On voit que $\lambda = 1$ est une valeur propre triple de la matrice A . Après échelonnage de $A - I$, on voit que celle-ci est de rang 2. Donc elle admet un seul vecteur propre \mathbf{u}_1 :

$$A - I \sim \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

On a donc une 1ère solution :

$$\mathbf{y}_1(x) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} e^x.$$

On construit un 1er vecteur propre généralisé solution du système

$$(A - I)\mathbf{u}_2 = \mathbf{u}_1.$$

Alors

$$\mathbf{u}_2 = \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix}$$

et

$$\mathbf{y}_2(x) = (x\mathbf{u}_1 + \mathbf{u}_2) e^x = \left(x \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} \right) e^x$$

est une 2ème solution indépendante. On construit un 2ème vecteur propre généralisé solution du système

$$(A - I)\mathbf{u}_3 = \mathbf{u}_2.$$

Alors,

$$\mathbf{u}_3 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

et

$$\mathbf{y}_3(x) = \left(\frac{x^2}{2} \mathbf{u}_1 + x\mathbf{u}_2 + \mathbf{u}_3 \right) e^x = \left(\frac{x^2}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \right) e^x$$

est une 3ème solution indépendante. \square

À l'exemple précédent, l'espace invariant associé à la valeur propre triple est de dimension 1. Donc la construction des deux vecteurs propres généralisés est toute simple. À l'exemple suivant, l'espace invariant associé à la valeur propre triple est de dimension 2. Donc la construction d'un vecteur propre généralisé est un peu plus complexe.

EXEMPLE 4.7. Résoudre le système $\mathbf{y}' = A\mathbf{y}$:

$$\mathbf{y}' = \begin{bmatrix} 1 & 2 & 1 \\ -4 & 7 & 2 \\ 4 & -4 & 1 \end{bmatrix} \mathbf{y}.$$

RÉSOLUTION. **(a) La solution analytique.**— On voit que $\lambda = 3$ est une valeur propre triple de la matrice A . Après échelonnage de $A - 3I$, on voit que celle-ci est de rang 1. Donc elle admet deux vecteurs propres indépendants \mathbf{u}_1 et \mathbf{u}_2 :

$$A - 3I \sim \begin{bmatrix} -2 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

On a donc deux solutions indépendantes :

$$\mathbf{y}_1(x) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} e^{3x}, \quad \mathbf{y}_2(x) = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} e^{3x}.$$

Pour obtenir une 3^{ème} solution indépendante on construit un vecteur propre généralisé solution du système

$$(A - 3I)\mathbf{u}_3 = \alpha\mathbf{u}_1 + \beta\mathbf{u}_2.$$

On choisit les paramètres α et β pour que le second membre

$$\mathbf{u}_4 = \alpha\mathbf{u}_1 + \beta\mathbf{u}_2 = \begin{bmatrix} \alpha + \beta \\ \alpha \\ 2\beta \end{bmatrix}$$

soit dans l'enveloppe linéaire V des colonnes de la matrice $(A - 3I)$. Puisque $\text{rang}(A - 3I) = 1$, alors

$$V = \text{lin} \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

et l'on peut prendre

$$\begin{bmatrix} \alpha + \beta \\ \alpha \\ 2\beta \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}.$$

Donc, $\alpha = 2$ et $\beta = -1$, d'où

$$\mathbf{u}_4 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

et

$$\mathbf{y}_3(x) = (x\mathbf{u}_4 + \mathbf{u}_3) e^{3x} = \left(x \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) e^{3x}$$

est une 3^{ème} solution indépendante.

(b) La solution Matlab symbolique.— La résolution symbolique utilise la forme normale de Jordan : $J = X^{-1}AX$, de la matrice A . On pose

$$\mathbf{y} = X\mathbf{w}$$

pour simplifier l'équation

$$\mathbf{w}' = J\mathbf{w}.$$

$$A = \begin{bmatrix} 1 & 2 & 1 \\ -4 & 7 & 2 \\ 4 & -4 & 1 \end{bmatrix}$$

$$[X, J] = \text{jordan}(A)$$

$$X = \begin{bmatrix} -2.0000 & 1.5000 & 0.5000 \\ -4.0000 & 0 & 0 \\ 4.0000 & 1.0000 & 1.0000 \end{bmatrix}$$

$$J = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

La matrice $J - 3I$ deux vecteurs propres :

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

et le vecteur propre généralisé :

$$\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

solution de l'équation

$$(J - 3I)\mathbf{u}_2 = \mathbf{u}_1.$$

On obtient donc trois solutions indépendantes :

$$\mathbf{y}_1 = e^{3x} X \mathbf{u}_1, \quad \mathbf{y}_2 = e^{3x} X(x\mathbf{u}_1 + \mathbf{u}_2), \quad \mathbf{y}_3 = e^{3x} X \mathbf{u}_3,$$

c'est-à-dire

$$\mathbf{u}_1 = [1 \ 0 \ 0]'; \quad \mathbf{u}_2 = [0 \ 1 \ 0]'; \quad \mathbf{u}_3 = [0 \ 0 \ 1]';$$

$$\text{syms } x; \quad \mathbf{y}_1 = \exp(3*x) * X * \mathbf{u}_1$$

$$\mathbf{y}_1 = \begin{bmatrix} -2*\exp(3*x) \\ -4*\exp(3*x) \\ 4*\exp(3*x) \end{bmatrix}$$

$$\mathbf{y}_2 = \exp(3*x) * X * (x*\mathbf{u}_1 + \mathbf{u}_2)$$

$$\mathbf{y}_2 = \begin{bmatrix} -2*\exp(3*x)*x + 3/2*\exp(3*x) \\ -4*\exp(3*x)*x \\ 4*\exp(3*x)*x + \exp(3*x) \end{bmatrix}$$

$$\mathbf{y}_3 = \exp(3*x) * X * \mathbf{u}_3$$

$$\mathbf{y}_3 = \begin{bmatrix} 1/2*\exp(3*x) \\ 0 \end{bmatrix}$$

[$\exp(3*x)$]

□

4.5. Systèmes linéaires non homogènes

Au chapitre 3, on a employé les méthodes des coefficients indéterminés et de la variation des paramètres pour trouver des solutions particulières d'équations différentielles non homogènes. Ici on généralise ces méthodes aux systèmes linéaires de la forme

$$\mathbf{y}' = A\mathbf{y} + \mathbf{f}(x). \quad (4.17)$$

On rappelle que la solution générale d'un système non homogène est la somme de la solution générale \mathbf{y}_h du système homogène

$$\mathbf{y}' = A\mathbf{y}$$

et d'une solution particulière \mathbf{y}_p de (4.17).

4.5.1. Méthode des coefficients indéterminés. On peut employer la méthode des coefficients indéterminés si la matrice A de (4.17) est constante et la dimension de l'espace vectoriel engendré par les dérivées du second membre $\mathbf{f}(x)$ de (4.17) est finie. C'est le cas quand $\mathbf{f}(x)$ est combinaison de cosinus, sinus, exponentielles, cosinus et sinus hyperboliques et polynômes. Pour ces problèmes un choix approprié de \mathbf{y}_p est une combinaison linéaire du vecteur $\mathbf{f}(x)$ et de ses dérivées indépendantes.

EXEMPLE 4.8. Trouver la solution générale du système linéaire non homogène

$$\mathbf{y}' = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{y} + \begin{bmatrix} 4e^{-3x} \\ e^{-2x} \end{bmatrix} := A\mathbf{y} + \mathbf{f}(x).$$

RÉSOLUTION. Les valeurs propres de la matrice A du système sont

$$\lambda_1 = i, \quad \lambda_2 = -i,$$

et les vecteurs propres associés sont

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

Donc la solution générale \mathbf{y}_h du système homogène est

$$\mathbf{y}_h(x) = k_1 e^{ix} \begin{bmatrix} 1 \\ i \end{bmatrix} + k_2 e^{-ix} \begin{bmatrix} 1 \\ -i \end{bmatrix},$$

où k_1 et k_2 sont des constantes complexes. Pour obtenir deux solutions réelles indépendantes on utilise le fait que la partie réelle et la partie imaginaire d'une solution d'une équation linéaire réelle sont aussi des solutions. On voit que la partie réelle et la partie imaginaire de la 1ère solution :

$$\begin{aligned} \mathbf{u}_1 &= \begin{bmatrix} 1 \\ i \end{bmatrix} e^{ix} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + i \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) (\cos x + i \sin x) \\ &= \begin{bmatrix} \cos x \\ -\sin x \end{bmatrix} + i \begin{bmatrix} \sin x \\ \cos x \end{bmatrix}, \end{aligned}$$

sont des solutions indépendantes. On obtient donc la solution générale réelle du système homogène de la forme

$$\mathbf{y}_h(x) = c_1 \begin{bmatrix} \cos x \\ -\sin x \end{bmatrix} + c_2 \begin{bmatrix} \sin x \\ \cos x \end{bmatrix}.$$

On écrit le second membre $\mathbf{f}(x)$ sous la forme

$$\mathbf{f}(x) = 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{-3x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-2x} := 4\mathbf{e}_1 e^{-3x} + \mathbf{e}_2 e^{-2x}$$

qui définit les vecteurs \mathbf{e}_1 et \mathbf{e}_2 . On voit que $\mathbf{f}(x)$ et $\mathbf{y}_h(x)$ n'ont aucune partie commune. Donc on prend $\mathbf{y}_p(x)$ de la forme

$$\mathbf{y}_p(x) = \mathbf{a} e^{-3x} + \mathbf{b} e^{-2x}.$$

Si l'on substitue $\mathbf{y}_p(x)$ dans le système donné on obtient

$$\mathbf{0} = (3\mathbf{a} + A\mathbf{a} + 4\mathbf{e}_1) e^{-3x} + (2\mathbf{b} + A\mathbf{b} + \mathbf{e}_2) e^{-2x}.$$

Puisque les fonctions e^{-3x} et e^{-2x} sont linéairement indépendantes, leurs coefficients sont nuls. On obtient donc deux équations pour \mathbf{a} et \mathbf{b} :

$$(A + 3I)\mathbf{a} = -4\mathbf{e}_1, \quad (A + 2I)\mathbf{b} = -\mathbf{e}_2.$$

Alors,

$$\mathbf{a} = -4(A + 3I)^{-1}\mathbf{e}_1 = -\frac{1}{5} \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

$$\mathbf{b} = -(A + 2I)^{-1}\mathbf{e}_2 = \frac{1}{5} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

et

$$\mathbf{y}_p(x) = - \begin{bmatrix} \frac{6}{5} e^{-3x} - \frac{1}{5} e^{-2x} \\ \frac{2}{5} e^{-3x} + \frac{2}{5} e^{-2x} \end{bmatrix}.$$

La solution générale est

$$\mathbf{y}(x) = \mathbf{y}_h(x) + \mathbf{y}_p(x). \quad \square$$

4.5.2. Méthode de la variation des paramètres. La méthode de la variation des paramètres s'applique, au moins en théorie, aux systèmes non homogènes à coefficients matriciels variables $A(x)$ et à second membre général $\mathbf{f}(x)$. Une matrice solution fondamentale

$$Y(x) = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n],$$

du système homogène

$$\mathbf{y}' = A(x)\mathbf{y}$$

est solution de l'équation

$$Y'(x) = A(x)Y(x).$$

Puisque les colonnes de $Y(x)$ sont linéairement indépendantes, la solution générale $\mathbf{y}_h(x)$ du système homogène est une combinaison linéaire de ses colonnes,

$$\mathbf{y}_h(x) = Y(x)\mathbf{c},$$

où \mathbf{c} est un n -vecteur arbitraire. La méthode de la variation des paramètres construit une solution particulière du système non homogène

$$\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$$

de la forme

$$\mathbf{y}_p(x) = Y(x)\mathbf{c}(x).$$

Si l'on substitue cette expression dans le système non homogène, on obtient

$$Y'\mathbf{c} + Y\mathbf{c}' = AY\mathbf{c} + \mathbf{f}.$$

Puisque $Y' = AY$, alors $Y'\mathbf{c} = AY\mathbf{c}$. Donc, l'expression antérieure se simplifie :

$$Y\mathbf{c}' = \mathbf{f}.$$

Puisque la matrice solution fondamentale est inversible, on a

$$\mathbf{c}'(x) = Y^{-1}(x)\mathbf{f}(x), \quad \text{ou} \quad \mathbf{c}(x) = \int^x Y^{-1}(s)\mathbf{f}(s) ds.$$

Il suit que

$$\mathbf{y}_p(x) = Y(x) \int^x Y^{-1}(s)\mathbf{f}(s) ds.$$

Pour un problème à valeur initiale

$$\mathbf{y}(x_0) = \mathbf{y}_0,$$

l'unique solution est

$$\mathbf{y}(x) = Y(x)Y^{-1}(x_0)\mathbf{y}_0 + Y(x) \int_{x_0}^x Y^{-1}(s)\mathbf{f}(s) ds.$$

Le lecteur est invité à résoudre l'exemple 4.8 par variation des paramètres.

Transformation de Laplace

5.1. Définition

DÉFINITION 5.1. Soit $f(t)$ une fonction définie sur $[0, +\infty[$. La transformée de Laplace $F(s)$ de la fonction $f(t)$ est définie par l'intégrale

$$\mathcal{L}(f)(s) := F(s) = \int_0^{\infty} e^{-st} f(t) dt, \quad (5.1)$$

sous la condition que l'intégrale existe pour $s > \gamma$. Dans ce cas on dit que $f(t)$ est transformable et qu'elle est l'*originale* de $F(s)$.

On voit que la fonction exponentielle

$$f(t) = e^{t^2}$$

n'admet pas de transformée de Laplace puisque l'intégrale (5.1) n'existe pas quelque soit le choix de $s > 0$.

On illustre la définition par quelques exemples utiles.

EXEMPLE 5.1. Trouver la transformée de Laplace de $f(t) = 1$.

RÉSOLUTION. (a) **Résolution analytique.**—

$$\begin{aligned} \mathcal{L}(1)(s) &= \int_0^{\infty} e^{-st} dt, \quad s > 0, \\ &= -\frac{1}{s} e^{-st} \Big|_0^{\infty} = -\frac{1}{s} (0 - 1) \\ &= \frac{1}{s}. \end{aligned}$$

(b) **Résolution par Matlab symbolique.**—

```
>> f = sym('Heaviside(t)');
>> F = laplace(f)
F = 1/s
```

La fonction `Heaviside` est une fonction de Maple. L'assistance pour les fonctions de Maple s'obtient par la commande `mhelp`. \square

EXEMPLE 5.2. Montrer que

$$\mathcal{L}(e^{at})(s) = \frac{1}{s-a}. \quad (5.2)$$

RÉSOLUTION. (a) Résolution analytique.—

$$\begin{aligned}\mathcal{L}(e^{at})(s) &= \int_0^{\infty} e^{-st} e^{at} dt \\ &= \int_0^{\infty} e^{-(s-a)t} dt \quad (s > a) \\ &= -\frac{1}{s-a} \left[e^{-(s-a)t} \right]_0^{\infty} \\ &= \frac{1}{s-a}.\end{aligned}$$

(b) Résolution par Matlab symbolique.—

```
>> syms a t;
>> f = exp(a*t);
>> F = laplace(f)
F = 1/(s-a)
```

□

THÉORÈME 5.1. *La transformation de Laplace,*

$$\mathcal{L} : f(t) \mapsto F(s),$$

est linéaire.

DÉMONSTRATION.

$$\begin{aligned}\mathcal{L}(af + bg) &= \int_0^{\infty} e^{-st} [af(t) + bg(t)] dt \\ &= a \int_0^{\infty} e^{-st} f(t) dt + b \int_0^{\infty} e^{-st} g(t) dt \\ &= a\mathcal{L}(f)(s) + b\mathcal{L}(g)(s). \quad \square\end{aligned}$$

EXEMPLE 5.3. Trouver la transformée de Laplace de $f(t) = \cosh at$.

RÉSOLUTION. (a) Résolution analytique.— Puisque

$$\cosh at = \frac{1}{2} (e^{at} + e^{-at}),$$

on obtient

$$\begin{aligned}\mathcal{L}(\cosh at)(s) &= \frac{1}{2} [\mathcal{L}(e^{at}) + \mathcal{L}(e^{-at})] \\ &= \frac{1}{2} \left[\frac{1}{s-a} + \frac{1}{s+a} \right] \\ &= \frac{s}{s^2 - a^2}.\end{aligned}$$

(b) Résolution par Matlab symbolique.—

```
>> syms a t;
>> f = cosh(a*t);
>> F = laplace(f)
F = s/(s^2-a^2)
```

□

EXEMPLE 5.4. Trouver la transformée de Laplace de $f(t) = \sinh at$.

RÉSOLUTION. (a) **Résolution analytique.**— Puisque

$$\sinh at = \frac{1}{2} (e^{at} - e^{-at}),$$

on obtient

$$\begin{aligned} \mathcal{L}(\sinh at)(s) &= \frac{1}{2} [\mathcal{L}(e^{at}) - \mathcal{L}(e^{-at})] \\ &= \frac{1}{2} \left[\frac{1}{s-a} - \frac{1}{s+a} \right] \\ &= \frac{a}{s^2 - a^2}. \end{aligned}$$

(b) **Résolution par Matlab symbolique.**—

```
>> syms a t;
>> f = sinh(a*t);
>> F = laplace(f)
F = a/(s^2-a^2)
```

□

REMARQUE 5.1. On voit que $\mathcal{L}(\cosh at)(s)$ est une fonction paire de a et $\mathcal{L}(\sinh at)(s)$ est une fonction impaire de a .

EXEMPLE 5.5. Trouver la transformée de Laplace de $f(t) = t^n$.

RÉSOLUTION. On procède par induction mathématique. Supposons que

$$\mathcal{L}(t^{n-1})(s) = \frac{(n-1)!}{s^n}.$$

Cette formule est vraie pour $n = 1$,

$$\mathcal{L}(1)(s) = \frac{0!}{s^1} = \frac{1}{s}.$$

Si $s > 0$ et $n \geq 2$, il suit par intégration par parties :

$$\begin{aligned} \mathcal{L}(t^n)(s) &= \int_0^\infty e^{-st} t^n dt \\ &= -\frac{1}{s} [t^n e^{-st}]_0^\infty + \frac{n}{s} \int_0^\infty e^{-st} t^{n-1} dt \\ &= \frac{n}{s} \mathcal{L}(t^{n-1})(s). \end{aligned}$$

Alors par récurrence:

$$\begin{aligned} \mathcal{L}(t^n)(s) &= \frac{n}{s} \frac{(n-1)!}{s^n} \\ &= \frac{n!}{s^{n+1}}, \quad s > 0. \quad \square \end{aligned}$$

On trouve la transformée de Laplace, par exemple, de t^5 , par les commandes de Matlab symbolique :

```

>> syms t
>> f = t^5;
>> F = laplace(f)
F = 120/s^6
ou
>> F = laplace(sym('t^5'))
F = 120/s^6
ou
>> F = laplace(sym('t')^5)
F = 120/s^6

```

EXEMPLE 5.6. Trouver la transformée de Laplace de $\cos \omega t$ et de $\sin \omega t$.

RÉSOLUTION. **(a) Résolution analytique.**— On emploie l'identité d'Euler :

$$e^{i\omega t} = \cos \omega t + i \sin \omega t, \quad i = \sqrt{-1}.$$

Alors

$$\begin{aligned}
 \mathcal{L}(e^{i\omega t})(s) &= \int_0^{\infty} e^{-st} e^{i\omega t} dt \quad (s > 0) \\
 &= \int_0^{\infty} e^{-(s-i\omega)t} dt \\
 &= -\frac{1}{s-i\omega} \left[e^{-(s-i\omega)t} \right]_0^{\infty} \\
 &= -\frac{1}{s-i\omega} [e^{-st} e^{i\omega t}]_{t \rightarrow \infty} - 1] \\
 &= \frac{1}{s-i\omega} = \frac{1}{s-i\omega} \frac{s+i\omega}{s+i\omega} \\
 &= \frac{s+i\omega}{s^2+\omega^2}.
 \end{aligned}$$

Par la linéarité de \mathcal{L} , on obtient

$$\begin{aligned}
 \mathcal{L}(e^{i\omega t})(s) &= \mathcal{L}(\cos \omega t + i \sin \omega t) \\
 &= \mathcal{L}(\cos \omega t) + i \mathcal{L}(\sin \omega t) \\
 &= \frac{s}{s^2+\omega^2} + i \frac{\omega}{s^2+\omega^2}
 \end{aligned}$$

On a donc

$$\mathcal{L}(\cos \omega t) = \frac{s}{s^2 + \omega^2}, \quad (5.3)$$

qui est paire en ω , et

$$\mathcal{L}(\sin \omega t) = \frac{\omega}{s^2 + \omega^2}, \quad (5.4)$$

qui est impaire en ω .

(b) Résolution par Matlab symbolique.—

```

>> syms omega t;
>> f = cos(omega*t);
>> g = sin(omega*t);
>> F = laplace(f)
F = s/(s^2+omega^2)

```

>> G = laplace(g)
 G = omega/(s^2+omega^2)

□

On supposera, dans la suite, que la transformée de Laplace des fonctions considérées dans ce chapitre existe et est différentiable et intégrable sous des conditions additionnelles. Les hypothèses de base sont énoncées dans la définition et le théorème suivants. La formule générale de la transformation de Laplace réciproque, qui n'est pas introduite dans ce chapitre, requiert aussi les résultats suivants.

DÉFINITION 5.2. On dit que la fonction $f(t)$ est de type exponentiel d'ordre γ s'il existe des constantes γ , $M > 0$ et $T > 0$ telles que

$$|f(t)| \leq M e^{\gamma t}, \quad \text{pour tout } t > T. \quad (5.5)$$

On appelle *abscisse de convergence* de $f(t)$ la borne inférieure $\gamma_0 \leq \gamma$ telle que (5.5) est valide.

THÉORÈME 5.2. Soit $f(t)$ une fonction continue par morceaux sur l'intervalle $[0, \infty)$. Si γ_0 est l'abscisse de convergence de $f(t)$, alors l'intégrale

$$\int_0^{\infty} e^{-st} f(t) dt$$

est absolument et uniformément convergente pour tout $s > \gamma_0$.

PROOF. On démontre la convergence absolue :

$$\begin{aligned} \left| \int_0^{\infty} e^{-st} f(t) dt \right| &\leq \int_0^{\infty} M e^{-(s-\gamma_0)t} dt \\ &= -\frac{M}{s-\gamma_0} e^{-(s-\gamma_0)t} \Big|_0^{\infty} = \frac{M}{s-\gamma_0}. \quad \square \end{aligned}$$

5.2. Transformées de dérivées et d'intégrales

En vue des applications aux équations différentielles, il faut savoir transformer la dérivée d'une fonction.

THÉORÈME 5.3.

$$\mathcal{L}(f')(s) = s\mathcal{L}(f) - f(0), \quad s > 0. \quad (5.6)$$

DÉMONSTRATION. On intègre par partie :

$$\begin{aligned} \mathcal{L}(f')(s) &= \int_0^{\infty} e^{-st} f'(t) dt \\ &= e^{-st} f(t) \Big|_0^{\infty} - (-s) \int_0^{\infty} e^{-st} f(t) dt \\ &= s\mathcal{L}(f)(s) - f(0). \quad \square \end{aligned}$$

REMARQUE 5.2. On montre par récurrence les formules suivantes :

$$\mathcal{L}(f'')(s) = s^2\mathcal{L}(f) - sf(0) - f'(0), \quad (5.7)$$

$$\mathcal{L}(f''')(s) = s^3\mathcal{L}(f) - s^2f(0) - sf'(0) - f''(0). \quad (5.8)$$

En effet,

$$\begin{aligned}\mathcal{L}(f'')(s) &= s\mathcal{L}(f')(s) - f'(0) \\ &= s[s\mathcal{L}(f)(s) - f(0)] - f'(0) \\ &= s^2\mathcal{L}(f) - sf(0) - f'(0)\end{aligned}$$

et

$$\begin{aligned}\mathcal{L}(f''')(s) &= s\mathcal{L}(f'')(s) - f''(0) \\ &= s[s^2\mathcal{L}(f) - sf(0) - f'(0)] - f''(0) \\ &= s^3\mathcal{L}(f) - s^2f(0) - sf'(0) - f''(0). \quad \square\end{aligned}$$

On démontre le théorème général suivant par récurrence.

THÉORÈME 5.4. *Soit $f(t), f'(t), \dots, f^{(n-1)}(t)$ continues pour $t \geq 0$ et $f^{(n)}(t)$ transformable pour $s \geq \gamma$. Alors*

$$\mathcal{L}(f^{(n)})(s) = s^n\mathcal{L}(f) - s^{n-1}f(0) - s^{n-2}f'(0) - \dots - f^{(n-1)}(0). \quad (5.9)$$

DÉMONSTRATION. La démonstration est par récurrence comme à la remarque 5.2 pour $n = 2$ et $n = 3$. \square

EXEMPLE 5.7. Résoudre le problème aux valeurs initiales :

$$y'' + 4y' + 3y = 0, \quad y(0) = 3, \quad y'(0) = 1.$$

Tracer la solution.

RÉSOLUTION. (a) Résolution analytique.— On note

$$\mathcal{L}(y)(s) = Y(s)$$

et l'on transforme l'équation différentielle :

$$\begin{aligned}\mathcal{L}(y'') + 4\mathcal{L}(y') + 3\mathcal{L}(y) &= s^2Y(s) - sy(0) - y'(0) + 4[sY(s) - y(0)] + 3Y(s) \\ &= 0.\end{aligned}$$

On obtient donc

$$(s^2 + 4s + 3)Y(s) - (s + 4)y(0) - y'(0) = 0,$$

qu'on récrit en remplaçant $y(0)$ et $y'(0)$ par leurs valeurs :

$$\begin{aligned}(s^2 + 4s + 3)Y(s) &= (s + 4)y(0) + y'(0) \\ &= 3(s + 4) + 1 = 3s + 13.\end{aligned}$$

On isole l'inconnue $Y(s)$ et l'on développe le 2ème membre en fractions simples :

$$\begin{aligned}Y(s) &= \frac{3s + 13}{s^2 + 4s + 3} \\ &= \frac{3s + 13}{(s + 1)(s + 3)} \\ &= \frac{A}{s + 1} + \frac{B}{s + 3}.\end{aligned}$$

Pour évaluer A et B , on chasse les dénominateurs et l'on récrit les deux derniers membres sous la forme

$$\begin{aligned} 3s + 13 &= (s + 3)A + (s + 1)B \\ &= (A + B)s + (3A + B). \end{aligned}$$

Des 1er et 3ème membres on obtient le système linéaire :

$$\begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 3 \\ 13 \end{bmatrix} \implies \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}.$$

Dans ce cas simple, on peut évaluer A et B tout simplement en posant $s = -1$ puis $s = -3$ dans l'identité

$$3s + 13 = (s + 3)A + (s + 1)B.$$

Alors

$$-3 + 13 = 2A \implies A = 5, \quad -9 + 13 = -2B \implies B = -2.$$

On a donc

$$Y(s) = \frac{5}{s+1} - \frac{2}{s+3}.$$

On trouve l'originale au moyen de la transformation de Laplace réciproque qui, en l'occurrence, est donnée de la formule (5.2) :

$$\begin{aligned} y(t) &= \mathcal{L}^{-1}(Y) = 5\mathcal{L}^{-1}\left(\frac{1}{s+1}\right) - 2\mathcal{L}^{-1}\left(\frac{1}{s+3}\right) \\ &= 5e^{-t} - 2e^{-3t}. \end{aligned}$$

(b) Résolution par Matlab symbolique.— On utilise l'expression pour $Y(s)$:

```
>> syms s t
>> Y = (3*s+13)/(s^2+4*s+3);
>> y = ilaplace(Y,s,t)
y = -2*exp(-3*t)+5*exp(-t)
```

(c) Résolution par Matlab numérique.— La fonction fichier M exp77.m :

```
function yprime = exp77(t,y);
yprime = [y(2); -3*y(1)-4*y(2)];
```

Le solveur ode45 de Matlab numérique produit la solution.

```
>> tspan = [0 4];
>> y0 = [3;1];
>> [t,y] = ode45('exp77',tspan,y0);
>> subplot(2,2,1); plot(t,y(:,1));
>> xlabel('t'); ylabel('y'); title('Graphe de la solution')
```

La commande subplot produit la fig. 5.1 dont le lettrage est suffisamment gros après réduction. \square

REMARQUE 5.3. On remarque que le polynôme caractéristique de l'équation différentielle homogène multiplie l'inconnue $Y(s)$ de l'équation transformée.

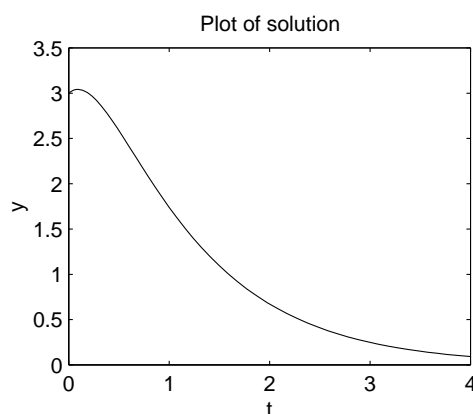


FIGURE 5.1. Graphe de la solution de l'équation différentielle de l'exemple 5.7.

REMARQUE 5.4. La méthode de résolution d'une équation différentielle homogène ou non homogène au moyen de la transformation de Laplace prend en main les valeurs initiales. Si l'équation est non homogène, ce procédé équivaut à la méthode des coefficients indéterminés ou de la variation des paramètres.

Puisque l'intégration est la réciproque de la dérivation et que la transformée de la dérivée de $f(t)$ est essentiellement la transformée de $f(t)$ multipliée par s on prévoit que la transformée de l'intégrale indéfinie de $f(t)$ sera la transformée de $f(t)$ divisée par s car la division est l'inverse de la multiplication.

THÉORÈME 5.5. Soit $f(t)$ transformable pour $s \geq \gamma$. Alors

$$\mathcal{L} \left\{ \int_0^t f(\tau) d\tau \right\} = \frac{1}{s} \mathcal{L}(f), \quad (5.10)$$

ou, en employant la transformation de Laplace inverse,

$$\mathcal{L}^{-1} \left\{ \frac{1}{s} F(s) \right\} = \int_0^t f(\tau) d\tau. \quad (5.11)$$

DÉMONSTRATION. Soit

$$g(t) = \int_0^t f(\tau) d\tau.$$

Alors, $g'(t) = f(t)$ et

$$\mathcal{L}(f(t)) = \mathcal{L}(g'(t)) = s\mathcal{L}(g(t)) - g(0).$$

Puisque $g(0) = 0$, on a $\mathcal{L}(f) = s\mathcal{L}(g)$, d'où (5.10). \square

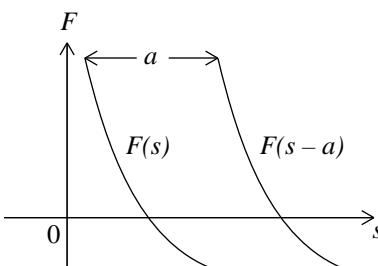
EXEMPLE 5.8. Soit

$$\mathcal{L}(f) = \frac{1}{s(s^2 + \omega^2)}.$$

Trouver $f(t)$.

RÉSOLUTION. (a) **Résolution analytique.**— Puisque

$$\mathcal{L}^{-1} \left(\frac{1}{s^2 + \omega^2} \right) = \frac{1}{\omega} \sin \omega t,$$

FIGURE 5.2. Déplacement $F(s-a)$ de $F(s)$ pour $a > 0$.

par (5.11) on a

$$\mathcal{L}^{-1} \left\{ \frac{1}{s} \left(\frac{1}{s^2 + \omega^2} \right) \right\} = \frac{1}{\omega} \int_0^t \sin \omega \tau \, d\tau = \frac{1}{\omega^2} (1 - \cos \omega t).$$

(b) Résolution par Matlab symbolique.—

```
>> syms s omega t
>> F = 1/(s*(s^2+omega^2));
>> f = ilaplace(F)
f = 1/omega^2-1/omega^2*cos(omega*t)
```

□

5.3. Déplacements en s et en t

Dans la pratique, on a besoin de l'originale de $F(s-a)$ et de la transformée de $u(t-a)f(t-a)$ où $u(t)$ est la fonction d'Heaviside,

$$u(t) = \begin{cases} 0, & \text{si } x < 0, \\ 1, & \text{si } x > 0. \end{cases} \quad (5.12)$$

THÉORÈME 5.6. *Soit*

$$\mathcal{L}(f)(s) = F(s), \quad s > \gamma.$$

Alors

$$\mathcal{L}(e^{at}f(t))(s) = F(s-a), \quad s-a > \gamma. \quad (5.13)$$

DÉMONSTRATION. (V. figure 5.2)

$$\begin{aligned} F(s-a) &= \int_0^{\infty} e^{-(s-a)t} f(t) \, dt \\ &= \int_0^{\infty} e^{-st} [e^{at} f(t)] \, dt \\ &= \mathcal{L}(e^{at}f(t))(s). \quad \square \end{aligned}$$

EXEMPLE 5.9. Appliquer le théorème 5.6 aux trois fonctions simples t^n , $\cos \omega t$ et $\sin \omega t$.

RÉSOLUTION. (a) Résolution analytique.— On présente les résultats sous forme de tableau.

| $f(t)$ | $F(s)$ | $e^{at}f(t)$ | $F(s-a)$ |
|-----------------|---------------------------------|------------------------|-------------------------------------|
| t^n | $\frac{n!}{s^{n+1}}$ | $e^{at}t^n$ | $\frac{n!}{(s-a)^{n+1}}$ |
| $\cos \omega t$ | $\frac{s}{s^2 + \omega^2}$ | $e^{at} \cos \omega t$ | $\frac{(s-a)}{(s-a)^2 + \omega^2}$ |
| $\sin \omega t$ | $\frac{\omega}{s^2 + \omega^2}$ | $e^{at} \sin \omega t$ | $\frac{\omega}{(s-a)^2 + \omega^2}$ |

(b) **Résolution par Matlab symbolique.**— Matlab symbolique produit la 2ème et la 3ème fonctions :

```
>> syms a t omega s;
>> f = exp(a*t)*cos(omega*t);
>> g = exp(a*t)*sin(omega*t);
>> F = laplace(f,t,s)
F = (s-a)/((s-a)^2+omega^2)
>> G = laplace(g,t,s)
G = omega/((s-a)^2+omega^2)
```

□

EXEMPLE 5.10. Trouver la solution du système amorti :

$$y'' + 2y' + 5y = 0, \quad y(0) = 2, \quad y'(0) = -4,$$

au moyen de la transformation de Laplace.

RÉSOLUTION. Posons

$$\mathcal{L}(y)(s) = Y(s).$$

Alors,

$$s^2Y(s) - sy(0) - y'(0) + 2[sY(s) - y(0)] + 5Y(s) = 0.$$

On regroupe $Y(s)$ au 1er membre :

$$\begin{aligned} (s^2 + 2s + 5)Y(s) &= sy(0) + y'(0) + 2y(0) \\ &= 2s - 4 + 4 \\ &= 2s. \end{aligned}$$

On isole $Y(s)$ et l'on réarrange le 2ème membre :

$$\begin{aligned} Y(s) &= \frac{2s}{s^2 + 2s + 1 + 4} \\ &= \frac{2(s+1) - 2}{(s+1)^2 + 2^2} \\ &= \frac{2(s+1)}{(s+1)^2 + 2^2} - \frac{2}{(s+1)^2 + 2^2}. \end{aligned}$$

On a donc la solution

$$y(t) = 2e^{-t} \cos 2t - e^{-t} \sin 2t. \quad \square$$

DÉFINITION 5.3. La translatée $u_a(t) = u(t-a)$ de la fonction d'Heaviside $u(t)$, dite échelon unité, est la fonction (V. figure 5.3) :

$$u_a(t) := u(t-a) = \begin{cases} 0, & \text{si } t < a, \\ 1, & \text{si } t > a, \end{cases} \quad a \geq 0. \quad (5.14)$$

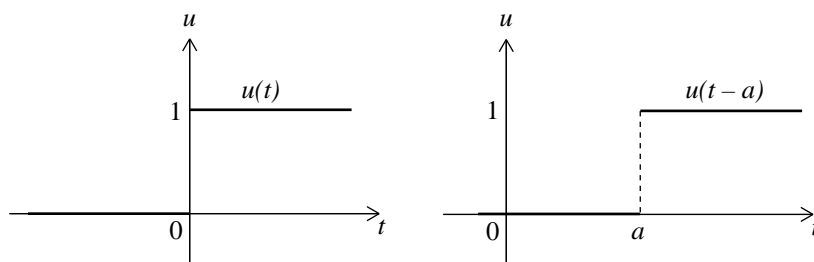


FIGURE 5.3. La fonction $u(t)$ d'Heaviside et sa tradatée $u(t - a)$, $a > 0$.

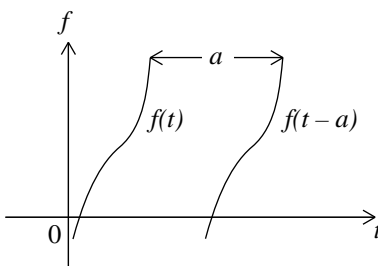


FIGURE 5.4. Déplacement $f(t - a)$ de $f(t)$ pour $a > 0$.

Selon les auteurs, on note $u(t)$ aussi par $\alpha(t)$ ou $H(t)$. En Matlab symbolique, la fonction d'Heaviside s'obtient par les commandes :

```
>> sym('Heaviside(t)')
>> u = sym('Heaviside(t)')
u = Heaviside(t)
```

L'assistance pour les fonctions de Maple s'obtient par la commande `mhelp`.

THÉORÈME 5.7. *Soit*

$$\mathcal{L}(f)(s) = F(s).$$

Alors

$$\mathcal{L}^{-1}(e^{-as}F(s)) = u(t - a)f(t - a), \quad (5.15)$$

c'est-à-dire

$$\mathcal{L}(u(t - a)f(t - a))(s) = e^{-as}F(s), \quad (5.16)$$

ou, par changement de variable en f on a la formule équivalente :

$$\mathcal{L}(u(t - a)f(t))(s) = e^{-as}\mathcal{L}(f(t + a))(s). \quad (5.17)$$

DÉMONSTRATION. (V. figure 5.4)

$$\begin{aligned}
e^{-as}F(s) &= e^{-as} \int_0^{\infty} e^{-s\tau} f(\tau) d\tau \\
&= \int_0^{\infty} e^{-s(\tau+a)} f(\tau) d\tau \\
&\quad (\text{On pose } \tau + a = t, d\tau = dt) \\
&= \int_a^{\infty} e^{-st} f(t-a) dt \\
&= \int_0^a e^{-st} 0 f(t-a) dt + \int_a^{\infty} e^{-st} 1 f(t-a) dt \\
&= \int_0^{\infty} e^{-st} u(t-a) f(t-a) dt \\
&= \mathcal{L}(u(t-a)f(t-a))(s). \quad \square
\end{aligned}$$

La formule équivalente (5.17) s'obtient par un changement de variable semblable :

$$\begin{aligned}
\mathcal{L}(u(t-a)f(t))(s) &= \int_0^{\infty} e^{-st} u(t-a) f(t) dt \\
&= \int_a^{\infty} e^{-st} f(t) dt \\
&\quad (\text{On pose } t = \tau + a, d\tau = dt) \\
&= \int_0^{\infty} e^{-s(\tau+a)} f(\tau+a) d\tau \\
&= e^{-as} \int_0^{\infty} e^{-s\tau} f(\tau+a) d\tau \\
&= e^{-as} \mathcal{L}(f(t+a))(s). \quad \square
\end{aligned}$$

La formule équivalente (5.17) peut simplifier les calculs comme on le verra dans certains exemples.

On voit que

$$\mathcal{L}(u(t-a)) = \frac{e^{-as}}{s}, \quad s > 0.$$

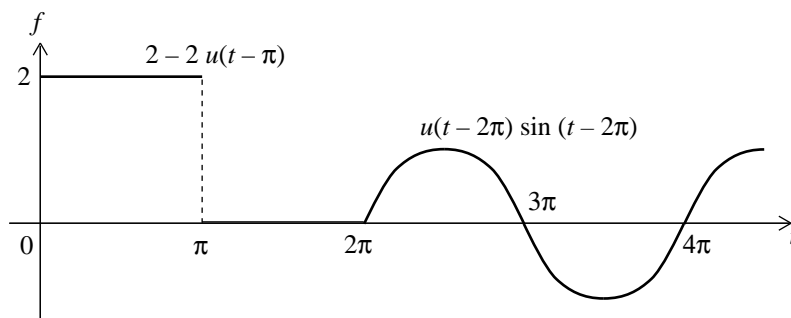
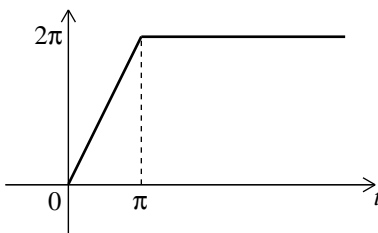
Cette formule découle directement de la définition :

$$\begin{aligned}
\mathcal{L}(u(t-a)) &= \int_0^{\infty} e^{-st} u(t-a) dt \\
&= \int_0^a e^{-st} 0 dt + \int_a^{\infty} e^{-st} 1 dt \\
&= -\frac{1}{s} e^{-st} \Big|_a^{\infty} = \frac{e^{-as}}{s}.
\end{aligned}$$

EXEMPLE 5.11. Soit

$$f(t) = \begin{cases} 2, & \text{si } 0 < t < \pi, \\ 0, & \text{si } \pi < t < 2\pi, \\ \sin t, & \text{si } 2\pi < t, \end{cases}$$

(V. figure 5.5). Trouver $F(s)$.

FIGURE 5.5. La fonction $f(t)$ de l'exemple 5.11.FIGURE 5.6. La fonction $f(t)$ de l'exemple 5.12.

RÉSOLUTION. On récrit $f(t)$ au moyen de la fonction d'Heaviside et de la 2π -périodicité de $\sin t$:

$$f(t) = 2 - 2u(t - \pi) + u(t - 2\pi) \sin(t - 2\pi).$$

Alors,

$$\begin{aligned} F(s) &= 2\mathcal{L}(1) - 2\mathcal{L}(u(t - \pi)1(t - \pi)) + \mathcal{L}(u(t - 2\pi) \sin(t - 2\pi)) \\ &= \frac{2}{s} - e^{-\pi s} \frac{2}{s} + e^{-2\pi s} \frac{1}{s^2 + 1}. \quad \square \end{aligned}$$

EXEMPLE 5.12. Soit

$$f(t) = \begin{cases} 2t, & \text{si } 0 < t < \pi, \\ 2\pi, & \text{si } \pi < t, \end{cases}$$

(V. figure 5.6). Trouver $F(s)$.

RÉSOLUTION. On récrit $f(t)$ au moyen de la fonction d'Heaviside :

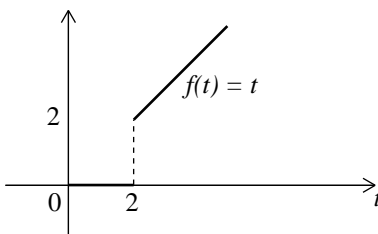
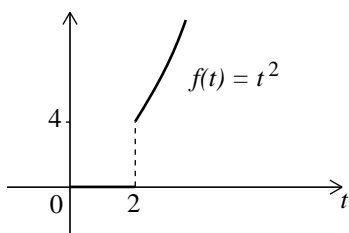
$$\begin{aligned} f(t) &= 2t - u(t - \pi)(2t) + u(t - \pi)2\pi \\ &= 2t - 2u(t - \pi)(t - \pi). \end{aligned}$$

Alors,

$$F(s) = \frac{2 \times 1!}{s^2} - 2e^{-\pi s} \frac{1}{s^2}. \quad \square$$

EXEMPLE 5.13. Soit

$$f(t) = \begin{cases} 0, & \text{si } 0 \leq t < 2, \\ t, & \text{si } 2 < t, \end{cases}$$

FIGURE 5.7. La fonction $f(t)$ de l'exemple 5.13.FIGURE 5.8. La fonction $f(t)$ de l'exemple 5.14.

(V. figure 5.7). Trouver $F(s)$.

RÉSOLUTION. On récrit $f(t)$ au moyen de la fonction d'Heaviside :

$$\begin{aligned} f(t) &= u(t-2)t \\ &= u(t-2)(t-2) + u(t-2)2. \end{aligned}$$

Alors, by (5.16) :

$$F(s) = e^{-2s} \frac{1!}{s^2} + 2 e^{-2s} \frac{0!}{s} = e^{-2s} \left[\frac{1}{s^2} + \frac{2}{s} \right]. \quad \square$$

De façon équivalente, par (5.17) :

$$\begin{aligned} \mathcal{L}(u(t-2)f(t))(s) &= e^{-2s} \mathcal{L}(f(t+2))(s) \\ &= e^{-2s} \mathcal{L}(t+2)(s) \\ &= e^{-2s} \left[\frac{1}{s^2} + \frac{2}{s} \right]. \quad \square \end{aligned}$$

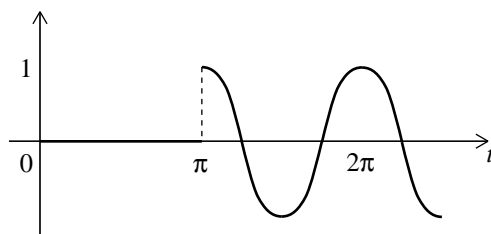
EXEMPLE 5.14. Soit

$$f(t) = \begin{cases} 0, & \text{si } 0 \leq t < 2, \\ t^2, & \text{si } 2 < t, \end{cases}$$

(V. figure 5.8). Trouver $F(s)$.

RÉSOLUTION. (a) **Résolution analytique.**— On récrit $f(t)$ au moyen de la fonction d'Heaviside :

$$\begin{aligned} f(t) &= u(t-2)t^2 \\ &= u(t-2)[(t-2)+2]^2 \\ &= u(t-2)[(t-2)^2 + 4(t-2) + 4]. \end{aligned}$$

FIGURE 5.9. La fonction $f(t)$ de l'exemple 5.15.

Alors, par (5.16) :

$$F(s) = e^{-2s} \left[\frac{2!}{s^3} + \frac{4}{s^2} + \frac{4}{s} \right].$$

De façon équivalente, par (5.17) :

$$\begin{aligned} \mathcal{L}(u(t-2)f(t))(s) &= e^{-2s} \mathcal{L}(f(t+2))(s) \\ &= e^{-2s} \mathcal{L}((t+2)^2)(s) \\ &= e^{-2s} \mathcal{L}(t^2 + 4t + 4)(s) \\ &= e^{-2s} \left[\frac{2!}{s^3} + \frac{4}{s^2} + \frac{4}{s} \right]. \end{aligned}$$

(b) Résolution par Matlab symbolique.—

```
syms s t
F = laplace('Heaviside(t-2)*((t-2)^2+4*(t-2)+4)')
F = 4*exp(-2*s)/s+4*exp(-2*s)/s^2+2*exp(-2*s)/s^3
```

□

EXEMPLE 5.15. Soit

$$F(s) = e^{-\pi s} \frac{s}{s^2 + 4}.$$

Trouver $f(t)$.

RÉSOLUTION. (a) Résolution analytique.— On voit que

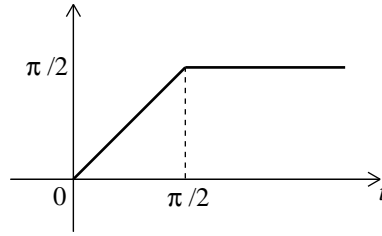
$$\begin{aligned} \mathcal{L}^{-1}F(s) &= u(t-\pi) \cos(2(t-\pi)) \\ &= \begin{cases} 0, & \text{si } 0 \leq t < \pi, \\ \cos(2(t-\pi)) = \cos 2t, & \text{si } \pi < t. \end{cases} \end{aligned}$$

On a tracé $f(t)$ dans la figure 5.9.

(b) Résolution par Matlab symbolique.—

```
>> syms s;
>> F = exp(-pi*s)*s/(s^2+4);
>> f = ilaplace(F)
f = Heaviside(t-pi)*cos(2*t)
```

□

FIGURE 5.10. La fonction $g(t)$ de l'exemple 5.16.

EXEMPLE 5.16. Résoudre le problème aux valeurs initiales :

$$y'' + 4y = g(t) = \begin{cases} t, & \text{si } 0 \leq t < \pi/2, \\ \pi/2, & \text{si } \pi/2 < t, \end{cases}$$

$$y(0) = 0, \quad y'(0) = 0,$$

au moyen de la transformation de Laplace.

RÉSOLUTION. Posons

$$\mathcal{L}(y) = Y(s) \quad \text{et} \quad \mathcal{L}(g) = G(s).$$

Alors,

$$\begin{aligned} \mathcal{L}(y'' + 4y) &= s^2 Y(s) - sy(0) - y'(0) + 4Y(s) \\ &= (s^2 + 4)Y(s) \\ &= G(s), \end{aligned}$$

où l'on a utilisé les valeurs données de $y(0)$ et $y'(0)$. On a donc

$$Y(s) = \frac{G(s)}{s^2 + 4}.$$

On récrit $g(t)$ (V. figure 5.10) au moyen de la fonction d'Heaviside :

$$\begin{aligned} g(t) &= t - u(t - \pi/2)t + u(t - \pi/2)\frac{\pi}{2} \\ &= t - u(t - \pi/2)(t - \pi/2), \end{aligned}$$

d'où

$$G(s) = \frac{1}{s^2} - e^{-(\pi/2)s} \frac{1}{s^2} = \left[1 - e^{-(\pi/2)s}\right] \frac{1}{s^2}.$$

Donc

$$Y(s) = \left[1 - e^{-(\pi/2)s}\right] \frac{1}{(s^2 + 4)s^2}.$$

On décompose la fonction rationnelle au 2ème terme en fractions simples :

$$\frac{1}{(s^2 + 4)s^2} = \frac{A}{s} + \frac{B}{s^2} + \frac{Cs + D}{s^2 + 4}.$$

On chasse les dénominateurs :

$$\begin{aligned} 1 &= (s^2 + 4)sA + (s^2 + 4)B + s^2(Cs + D) \\ &= (A + C)s^3 + (B + D)s^2 + 4As + 4B. \end{aligned}$$

On identifie les coefficients :

$$\begin{aligned} 4A &= 0 \implies A = 0, \\ 4B &= 1 \implies B = \frac{1}{4}, \\ B + D &= 0 \implies D = -\frac{1}{4}, \\ A + C &= 0 \implies C = 0, \end{aligned}$$

d'où

$$\frac{1}{(s^2 + 4)s^2} = \frac{1}{4} \frac{1}{s^2} - \frac{1}{4} \frac{1}{s^2 + 4}.$$

Donc

$$Y(s) = \frac{1}{4} \frac{1}{s^2} - \frac{1}{8} \frac{2}{s^2 + 2^2} - \frac{1}{4} e^{-(\pi/2)s} \frac{1}{s^2} + \frac{1}{8} e^{-(\pi/2)s} \frac{2}{s^2 + 2^2}$$

et, par la transformation de Laplace inverse,

$$y(t) = \frac{1}{4} t - \frac{1}{8} \sin 2t - \frac{1}{4} u(t - \pi/2)(t - \pi/2) + \frac{1}{8} u(t - \pi/2) \sin(2[t - \pi/2]). \quad \square$$

On peut aussi trouver l'originale de la transformée

$$Y(s) = \frac{1}{2} \left[1 - e^{-(\pi/2)s} \right] \frac{2}{(s^2 + 4)s^2}$$

de l'exemple 5.16 par deux intégrations au moyen de la formule (5.11) du théorème 5.5 :

$$\begin{aligned} \mathcal{L}^{-1} \left(\frac{1}{s} \frac{2}{s^2 + 2^2} \right) &= \int_0^t \sin 2\tau \, d\tau = \frac{1}{2} - \frac{1}{2} \cos(2t), \\ \mathcal{L}^{-1} \left(\frac{1}{s} \left[\frac{1}{s} \frac{2}{s^2 + 2^2} \right] \right) &= \frac{1}{2} \int_0^t (1 - \cos 2\tau) \, d\tau = \frac{t}{2} - \frac{1}{4} \sin(2t). \end{aligned}$$

On obtient alors l'originale $y(t)$ par la formule (5.15) du théorème 5.7.

5.4. La fonction delta de Dirac

Soit la fonction

$$f_k(t; a) = \begin{cases} 1/k, & \text{si } a \leq t \leq a + k, \\ 0, & \text{sinon.} \end{cases} \quad (5.18)$$

On voit que l'intégrale de $f_k(t; a)$ est 1 :

$$I_k = \int_0^\infty f_k(t; a) \, dt = \int_a^{a+k} \frac{1}{k} \, dt = 1. \quad (5.19)$$

On note

$$\delta(t - a)$$

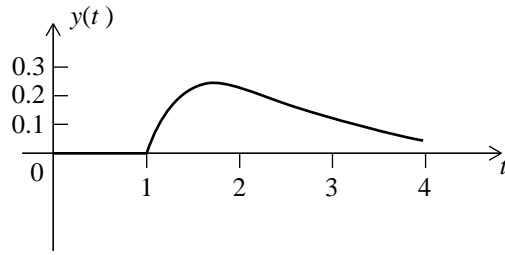
la limite de $f_k(t; a)$ quand $k \rightarrow 0$. Cette limite s'appelle *fonction delta de Dirac*, ou plus précisément, masse ou mesure de Dirac.

On peut représenter $f_k(t; a)$ au moyen de la différence de deux fonctions d'Heaviside :

$$f_k(t; a) = \frac{1}{k} [u(t - a) - u(t - (a + k))].$$

On obtient sa transformée de Laplace par (5.17),

$$\mathcal{L}(f_k(t; a)) = \frac{1}{ks} [e^{-as} - e^{-(a+k)s}] = e^{-as} \frac{1 - e^{-ks}}{ks}. \quad (5.20)$$

FIGURE 5.11. La solution $y(t)$ de l'exemple 5.17.

Le quotient au dernier membre tend vers 1 quand $k \rightarrow 0$ comme on peut voir par la règle de l'Hôpital. On a donc

$$\mathcal{L}(\delta(t-a)) = e^{-as}. \quad (5.21)$$

Matlab symbolique produit la transformée de Laplace de la fonction symbolique $\delta(t)$:

```
>> f = sym('Dirac(t)');
>> F = laplace(f)
F = 1
```

EXEMPLE 5.17. Résoudre le système amorti

$$y'' + 3y' + 2y = \delta(t-a), \quad y(0) = 0, \quad y'(0) = 0,$$

au repos pour $0 \leq t < a$ et percuté à l'instant $t = a$.

RÉSOLUTION. Par (5.21), la transformée de l'équation différentielle est

$$s^2Y + 3sY + 2Y = e^{-as}.$$

On isole $Y(s)$:

$$Y(s) = F(s)e^{-as},$$

où

$$F(s) = \frac{1}{(s+1)(s+2)} = \frac{1}{s+1} - \frac{1}{s+2}.$$

Alors,

$$f(t) = \mathcal{L}^{-1}(F) = e^{-t} - e^{-2t}.$$

Donc par (5.15),

$$\begin{aligned} y(t) &= \mathcal{L}^{-1}(e^{-as}F(s)) \\ &= f(t-a)u(t-a) \\ &= \begin{cases} 0, & \text{si } 0 \leq t < a, \\ e^{-(t-a)} - e^{-2(t-a)}, & \text{si } t > a. \end{cases} \end{aligned}$$

On illustre la solution dans la figure 5.11 pour $a = 1$. □

5.5. Dérivée et intégrale de la transformée

On démontre les formules suivantes.

THÉORÈME 5.8. *Si $F(s) = \mathcal{L}\{f(t)\}(s)$, alors*

$$\mathcal{L}\{tf(t)\}(s) = -F'(s), \quad (5.22)$$

ou, en se servant de la transformation de Laplace inverse,

$$\mathcal{L}^{-1}\{F'(s)\} = -tf(t). \quad (5.23)$$

De plus, si la limite

$$\lim_{t \rightarrow 0^+} \frac{f(t)}{t}$$

existe, alors

$$\mathcal{L}\left\{\frac{f(t)}{t}\right\}(s) = \int_s^\infty F(\tilde{s}) d\tilde{s}, \quad (5.24)$$

ou, en se servant de la transformation de Laplace inverse,

$$\mathcal{L}^{-1}\left\{\int_s^\infty F(\tilde{s}) d\tilde{s}\right\} = \frac{1}{t} f(t). \quad (5.25)$$

DÉMONSTRATION. Soit

$$F(s) = \int_0^\infty e^{-st} f(t) dt.$$

Par le théorème 5.2, (5.22) suit par dérivation :

$$\begin{aligned} F'(s) &= - \int_0^\infty e^{-st} [tf(t)] dt \\ &= -\mathcal{L}\{tf(t)\}(s). \end{aligned}$$

D'autre part, par le théorème 5.2, (5.24) suit par intégration :

$$\begin{aligned} \int_s^\infty F(\tilde{s}) d\tilde{s} &= \int_s^\infty \int_0^\infty e^{-\tilde{s}t} f(t) dt d\tilde{s} \\ &= \int_0^\infty f(t) \left[\int_s^\infty e^{-\tilde{s}t} d\tilde{s} \right] dt \\ &= \int_0^\infty f(t) \left[-\frac{1}{t} e^{-\tilde{s}t} \right]_{\tilde{s}=s}^{\tilde{s}=\infty} dt \\ &= \int_0^\infty e^{-st} \left[\frac{1}{t} f(t) \right] dt \\ &= \mathcal{L}\left\{\frac{1}{t} f(t)\right\}. \quad \square \end{aligned}$$

On généralise la formule (5.22) dans le théorème suivant.

THÉORÈME 5.9. *Si $t^n f(t)$ est transformable, alors*

$$\mathcal{L}\{t^n f(t)\}(s) = (-1)^n F^{(n)}(s), \quad (5.26)$$

ou, réciproquement,

$$\mathcal{L}^{-1}\{F^{(n)}(s)\} = (-1)^n t^n f(t). \quad (5.27)$$

EXEMPLE 5.18. Obtenir l'originale de $\frac{1}{(s+1)^2}$ par (5.23).

RÉSOLUTION. Posons

$$\frac{1}{(s+1)^2} = -\frac{d}{ds} \left(\frac{1}{s+1} \right) =: -F'(s).$$

Alors par (5.23)

$$\begin{aligned} \mathcal{L}^{-1}\{-F'(s)\} &= tf(t) = t \mathcal{L}^{-1} \left\{ \frac{1}{s+1} \right\} \\ &= t e^{-t}. \quad \square \end{aligned}$$

EXEMPLE 5.19. Etant donné $f(t)$ trouver $F(s)$ par (5.22).

$$\begin{array}{ll} f(t) & F(s) \\ \frac{1}{2\beta^3} [\sin \beta t - \beta t \cos \beta t] & \frac{1}{(s^2 + \beta^2)^2}, \end{array} \quad (5.28)$$

$$\begin{array}{ll} \frac{t}{2\beta} \sin \beta t & \frac{s}{(s^2 + \beta^2)^2}, \end{array} \quad (5.29)$$

$$\begin{array}{ll} \frac{1}{2\beta} [\sin \beta t + \beta t \cos \beta t] & \frac{s^2}{(s^2 + \beta^2)^2}. \end{array} \quad (5.30)$$

RÉSOLUTION. (a) **Résolution analytique.**— On applique (5.22) au 1er membre de (5.29) :

$$\begin{aligned} \mathcal{L}(t \sin \beta t)(s) &= -\frac{d}{ds} \left[\frac{\beta}{s^2 + \beta^2} \right] \\ &= \frac{2\beta s}{(s^2 + \beta^2)^2}, \end{aligned}$$

d'où, après division par 2β , on obtient le 2ème membre de (5.29).

De même, par (5.22), on obtient

$$\begin{aligned} \mathcal{L}(t \cos \beta t)(s) &= -\frac{d}{ds} \left[\frac{s}{s^2 + \beta^2} \right] \\ &= -\frac{s^2 + \beta^2 - 2s^2}{(s^2 + \beta^2)^2} \\ &= \frac{s^2 - \beta^2}{(s^2 + \beta^2)^2}. \end{aligned}$$

Alors

$$\begin{aligned} \mathcal{L} \left(t \cos \beta t \pm \frac{1}{\beta} \sin \beta t \right) (s) &= \frac{s^2 - \beta^2}{(s^2 + \beta^2)^2} \pm \frac{1}{s^2 + \beta^2} \\ &= \frac{s^2 - \beta^2 \pm (s^2 + \beta^2)}{(s^2 + \beta^2)^2}. \end{aligned}$$

Si l'on prend le signe + et l'on divise par deux, on obtient (5.30). Si l'on prend le signe - et l'on divise par $-2\beta^2$, on obtient (5.28).

(b) **Résolution par Matlab symbolique.**—

```
>> syms t beta s
>> f = (sin(beta*t)-beta*t*cos(beta*t))/(2*beta^3);
>> F = laplace(f,t,s)
F = 1/2/beta^3*(beta/(s^2+beta^2)-beta*(-1/(s^2+beta^2)+2*s^2/(s^2+beta^2)^2))
```

```

>> FF = simple(F)
FF = 1/(s^2+beta^2)^2

>> g = t*sin(beta*t)/(2*beta);
>> G = laplace(g,t,s)
G = 1/(s^2+beta^2)^2*s

>> h = (sin(beta*t)+beta*t*cos(beta*t))/(2*beta);
>> H = laplace(h,t,s)
H = 1/2/beta*(beta/(s^2+beta^2)+beta*(-1/(s^2+beta^2)+2*s^2/(s^2+beta^2)^2))
>> HH = simple(H)
HH = s^2/(s^2+beta^2)^2

```

□

EXEMPLE 5.20. Trouver

$$\mathcal{L}^{-1} \left\{ \ln \left(1 + \frac{\omega^2}{s^2} \right) \right\} (t).$$

RÉSOLUTION. (a) **Résolution analytique.**— On a

$$\begin{aligned}
 -\frac{d}{ds} \ln \left(1 + \frac{\omega^2}{s^2} \right) &= -\frac{d}{ds} \ln \left(\frac{s^2 + \omega^2}{s^2} \right) \\
 &= -\frac{s^2}{s^2 + \omega^2} \frac{2s^3 - 2s(s^2 + \omega^2)}{s^4} \\
 &= \frac{2\omega^2}{s(s^2 + \omega^2)} \\
 &= 2 \frac{(\omega^2 + s^2) - s^2}{s(s^2 + \omega^2)} \\
 &= \frac{2}{s} - 2 \frac{s}{s^2 + \omega^2} \\
 &=: F(s).
 \end{aligned}$$

Alors

$$f(t) = \mathcal{L}^{-1}(F) = 2 - 2 \cos \omega t.$$

Puisque

$$\frac{f(t)}{t} = 2\omega \frac{1 - \cos \omega t}{\omega t} \rightarrow 0 \quad \text{quand } t \rightarrow 0,$$

et

$$\begin{aligned}
 \int_s^\infty F(\tilde{s}) d\tilde{s} &= -\int_s^\infty \frac{d}{d\tilde{s}} \ln \left(1 + \frac{\omega^2}{\tilde{s}^2} \right) d\tilde{s} \\
 &= -\ln \left(1 + \frac{\omega^2}{\tilde{s}^2} \right) \Big|_s^\infty \\
 &= -\ln 1 + \ln \left(1 + \frac{\omega^2}{s^2} \right) \\
 &= \ln \left(1 + \frac{\omega^2}{s^2} \right),
 \end{aligned}$$

par (5.25) on a

$$\begin{aligned}\mathcal{L}^{-1}\left\{\ln\left(1+\frac{\omega^2}{s^2}\right)\right\} &= \mathcal{L}^{-1}\left\{\int_s^\infty F(\tilde{s})d\tilde{s}\right\} \\ &= \frac{1}{t}f(t) \\ &= \frac{2}{t}(1-\cos\omega t).\end{aligned}$$

(b) Résolution par Matlab symbolique.—

```
>> syms omega t s
>> F = log(1+(omega^2/s^2));
>> f = ilaplace(F,s,t)
f = 2/t-2/t*cos(omega*t)
```

□

5.6. Équation différentielle de Laguerre

On peut résoudre certaines équations différentielles à coefficients variables de la forme $at + b$ au moyen de la transformation de Laplace. En effet, par (5.22), (5.6) et (5.7), on obtient

$$\begin{aligned}\mathcal{L}(ty'(t)) &= -\frac{d}{ds}[sY(s) - y(0)] \\ &= -Y(s) - sY'(s),\end{aligned}\tag{5.31}$$

$$\begin{aligned}\mathcal{L}(ty''(t)) &= -\frac{d}{ds}[s^2Y(s) - sy(0) - y'(0)] \\ &= -2sY(s) - s^2Y'(s) + y(0).\end{aligned}\tag{5.32}$$

EXEMPLE 5.21. Trouver les solutions polynomiales $L_n(t)$ de l'équation de Laguerre :

$$ty'' + (1-t)y' + ny = 0, \quad n = 0, 1, \dots\tag{5.33}$$

RÉSOLUTION. La transformée de Laplace de l'équation (5.33) est

$$\begin{aligned}-2sY(s) - s^2Y'(s) + y(0) + sY(s) - y(0) + Y(s) + sY'(s) + nY(s) \\ = (s - s^2)Y'(s) + (n + 1 - s)Y(s) = 0.\end{aligned}$$

Cette équation est séparable:

$$\begin{aligned}\frac{dY}{Y} &= \frac{n+1-s}{(s-1)s} ds \\ &= \left(\frac{n}{s-1} - \frac{n+1}{s}\right) ds,\end{aligned}$$

d'où sa solution

$$\begin{aligned}\ln|Y(s)| &= n \ln|s-1| - (n+1) \ln s \\ &= \ln\left|\frac{(s-1)^n}{s^{n+1}}\right|,\end{aligned}$$

c'est-à-dire

$$Y(s) = \frac{(s-1)^n}{s^{n+1}}.$$

Il n'y a pas lieu d'ajouter une constante additive à la solution $\ln |Y(s)|$ parce qu'il en résulterait une constante multiplicative sur $Y(s)$ et cette dernière serait encore une solution puisque l'équation de Laguerre (5.33) est homogène.

Posons

$$L_n(t) = \mathcal{L}^{-1}(Y)(t),$$

où, par exception, la lettre majuscule L dans L_n désigne une fonction de t . En effet, on note $L_n(t)$ le polynôme de Laguerre de degré n . On montre que

$$L_0(t) = 1, \quad L_n(t) = \frac{e^t}{n!} \frac{d^n}{dt^n} (t^n e^{-t}), \quad n = 1, 2, \dots$$

On voit bien que $L_n(t)$ est un polynôme de degré n puisque les exponentielles s'annulent après la différentiation. Puisque, par le théorème 5.4,

$$\mathcal{L}(f^{(n)})(s) = s^n F(s) - s^{n-1} f(0) - s^{n-2} f'(0) - \dots - f^{(n-1)}(0),$$

on a

$$\mathcal{L} \left\{ (t^n e^{-t})^{(n)} \right\} (s) = s^n \frac{n!}{(s+1)^{n+1}},$$

et, par conséquent,

$$\begin{aligned} \mathcal{L} \left\{ \frac{e^t}{n!} (t^n e^{-t})^{(n)} \right\} &= \frac{n!}{n!} \frac{(s-1)^n}{s^{n+1}} \\ &= Y(s) \\ &= \mathcal{L}(L_n). \quad \square \end{aligned}$$

Les 4 premiers polynômes de Laguerre sont (V. figure 5.12) :

$$\begin{aligned} L_0(x) &= 1, & L_1(x) &= 1 - x, \\ L_2(x) &= 1 - 2x + \frac{1}{2}x^2, & L_3(x) &= 1 - 3x + \frac{3}{2}x^2 - \frac{1}{6}x^3. \end{aligned}$$

On peut obtenir les $L_n(x)$ par la récurrence:

$$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x).$$

Les polynômes de Laguerre satisfont les relations d'orthogonalité suivantes avec le poids $p(x) = e^{-x}$:

$$\int_0^\infty e^{-x} L_m(x) L_n(x) dx = \begin{cases} 0, & m \neq n, \\ 1, & m = n. \end{cases}$$

Matlab produit les polynômes de Laguerre:

```
>>L0 = dsolve('t*D2y+(1-t)*Dy=0','y(0)=1','t')
L0 = 1

>>L1 = dsolve('t*D2y+(1-t)*Dy+y=0','y(0)=1','t');
>> L1 = simple(L1)
L1 = 1-t

>> L2 = dsolve('t*D2y+(1-t)*Dy+2*y=0','y(0)=1','t');
>> L2 = simple(L2)
L2 = 1-2*t+1/2*t^2
```

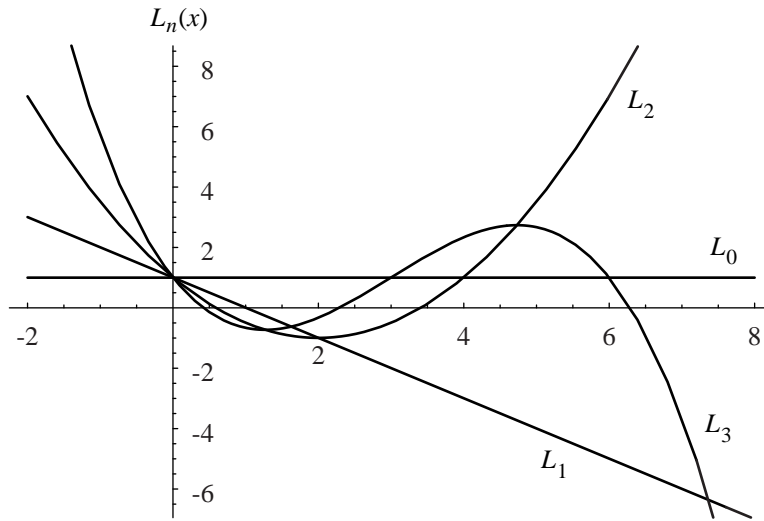


FIGURE 5.12. Les 4 premiers polynômes de LAGUERRE.

ainsi de suite. La commande `simple` de Matlab symbolique a pour but, nonorthodoxe en mathématiques, de simplifier une expression afin d'en réduire le nombre de caractères.

5.7. Convolution

L'originale du produit de deux transformées est la convolution des deux originales.

DÉFINITION 5.4. La convolution de $f(t)$ et de $g(t)$, notée $(f * g)(t)$, est la fonction

$$h(t) = \int_0^t f(\tau)g(t - \tau) d\tau. \quad (5.34)$$

On dit “ f convoluée avec g ”.

On vérifie que la convolution est commutative :

$$\begin{aligned} (f * g)(t) &= \int_0^t f(\tau)g(t - \tau) d\tau \\ &\quad (\text{on pose } t - \tau = \sigma, d\tau = -d\sigma) \\ &= - \int_t^0 f(t - \sigma)g(\sigma) d\sigma \\ &= \int_0^t g(\sigma)f(t - \sigma) d\sigma \\ &= (g * f)(t). \end{aligned}$$

THÉORÈME 5.10. *Soit*

$$F(s) = \mathcal{L}(f), \quad G(s) = \mathcal{L}(g), \quad H(s) = F(s)G(s), \quad h(t) = \mathcal{L}^{-1}(H).$$

Alors

$$h(t) = (f * g)(t) = \mathcal{L}^{-1}(F(s)G(s)). \quad (5.35)$$

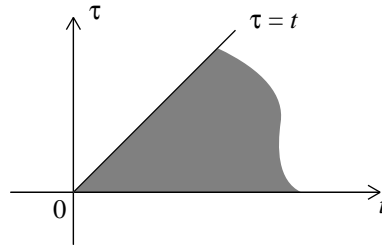


FIGURE 5.13. Région d'intégration du plan $t\tau$ utilisée dans la démonstration du théorème 5.10.

DÉMONSTRATION. Par définition et par (5.16), on a

$$\begin{aligned} e^{-s\tau}G(s) &= \mathcal{L}(g(t-\tau)u(t-\tau)) \\ &= \int_0^\infty e^{-st}g(t-\tau)u(t-\tau) dt \\ &= \int_\tau^\infty e^{-st}g(t-\tau) dt. \end{aligned}$$

Par ceci et par la définition de $F(s)$ on a

$$\begin{aligned} F(s)G(s) &= \int_0^\infty e^{-s\tau}f(\tau)G(s) d\tau \\ &= \int_0^\infty f(\tau) \left[\int_\tau^\infty e^{-st}g(t-\tau) dt \right] d\tau, \quad (s > \gamma) \\ &= \int_0^\infty e^{-st} \left[\int_0^t f(\tau)g(t-\tau) d\tau \right] dt \\ &= \mathcal{L}[(f * g)(t)](s) \\ &= \mathcal{L}(h)(s). \end{aligned}$$

On illustre dans la figure 5.13 le domaine d'intégration dans le plan $t\tau$ utilisé dans la démonstration du théorème 5.10. \square

EXEMPLE 5.22. Calculer $(1 * 1)(t)$.

RÉSOLUTION.

$$(1 * 1)(t) = \int_0^t 1 \times 1 d\tau = t. \quad \square$$

EXEMPLE 5.23. Calculer $e^t * e^t$.

RÉSOLUTION.

$$\begin{aligned} e^t * e^t &= \int_0^t e^\tau e^{t-\tau} d\tau \\ &= \int_0^t e^t d\tau = t e^t. \quad \square \end{aligned}$$

EXEMPLE 5.24. Calculer l'originale de

$$\frac{1}{(s-a)(s-b)}, \quad a \neq b,$$

au moyen de la convolution.

RÉSOLUTION.

$$\begin{aligned}
 \mathcal{L}^{-1} \left[\frac{1}{(s-a)(s-b)} \right] &= e^{at} * e^{bt} \\
 &= \int_0^t e^{a\tau} e^{b(t-\tau)} d\tau \\
 &= e^{bt} \int_0^t e^{(a-b)\tau} d\tau \\
 &= e^{bt} \frac{1}{a-b} e^{(a-b)\tau} \Big|_0^t \\
 &= \frac{e^{bt}}{a-b} [e^{(a-b)t} - 1] \\
 &= \frac{e^{at} - e^{bt}}{a-b}. \quad \square
 \end{aligned}$$

On peut résoudre certaines équations intégrales au moyen de la transformation de Laplace.

EXEMPLE 5.25. Résoudre l'équation intégrale

$$y(t) = t + \int_0^t y(\tau) \sin(t - \tau) d\tau. \quad (5.36)$$

RÉSOLUTION. Le dernier terme de (5.36) est une convolution. Alors

$$y(t) = t + y * \sin t.$$

Donc

$$Y(s) = \frac{1}{s^2} + Y(s) \frac{1}{s^2 + 1},$$

d'où

$$Y(s) = \frac{s^2 + 1}{s^4} = \frac{1}{s^2} + \frac{1}{s^4}.$$

Enfin,

$$y(t) = t + \frac{1}{6} t^3. \quad \square$$

5.8. Fractions simples

La décomposition d'une fonction rationnelle en fractions simples a été étudiée dans le cours de calcul différentiel et intégral.

Il suffit de mentionner que si $p(\lambda)$ est le polynôme caractéristique d'une équation différentielle $Ly = r(t)$ à coefficients constants, la mise en facteur de $p(\lambda)$ requise pour trouver les zéros de p et par conséquent les solutions indépendantes de $Ly = 0$, est aussi requise pour la réduction de $1/p(s)$ en fractions simples avec l'emploi de la transformation de Laplace.

Le phénomène de résonance correspond à des racines multiples.

Avec la boîte symbolique de Matlab, on peut obtenir le développement en fractions simples par la commande `convert`, elle-même référencée par la commande `mhelp convert[parfrac]`.

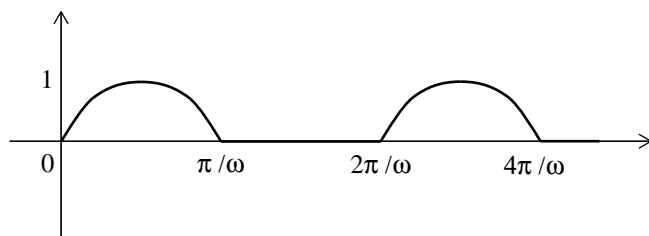


FIGURE 5.14. Semi-rectification de l'onde de l'exemple 5.26.

5.9. Transformées de fonctions périodiques

DÉFINITION 5.5. Une fonction $f(t)$ définie pour tout $t > 0$ est périodique de période p , $p > 0$, si

$$f(t + p) = f(t), \quad \text{pour tout } t > 0. \quad (5.37)$$

THÉORÈME 5.11. Soit $f(t)$ une fonction p -périodique. Alors

$$\mathcal{L}(f)(s) = \frac{1}{1 - e^{-ps}} \int_0^p e^{-st} f(t) dt, \quad s > 0. \quad (5.38)$$

DÉMONSTRATION. On exploite la périodicité de f :

$$\begin{aligned} \mathcal{L}(f)(s) &= \int_0^{\infty} e^{-st} f(t) dt \\ &= \int_0^p e^{-st} f(t) dt + \int_p^{2p} e^{-st} f(t) dt + \int_{2p}^{3p} e^{-st} f(t) dt + \dots \end{aligned}$$

On substitue

$$t = \tau + p, \quad t = \tau + 2p, \quad \dots,$$

respectivement dans la 2ème, la 3ème intégrales, etc. Les nouvelles limites d'intégration sont alors 0 et p . Donc

$$\begin{aligned} \mathcal{L}(f)(s) &= \int_0^p e^{-st} f(t) dt + \int_0^p e^{-s(t+p)} f(t) dt + \int_0^p e^{-s(t+2p)} f(t) dt + \dots \\ &= (1 + e^{-sp} + e^{-2sp} + \dots) \int_0^p e^{-st} f(t) dt \\ &= \frac{1}{1 - e^{-ps}} \int_0^p e^{-st} f(t) dt. \quad \square \end{aligned}$$

EXEMPLE 5.26. Trouver la transformée de Laplace de la semi-rectification de l'onde

$$f(t) = \sin \omega t$$

(V. figure 5.14).

RÉSOLUTION. (a) **Résolution analytique.**— L'onde semi-rectifiée de période $p = 2\pi/\omega$ est

$$f(t) = \begin{cases} \sin \omega t, & \text{si } 0 < t < \pi/\omega, \\ 0, & \text{si } \pi/\omega < t < 2\pi/\omega. \end{cases}$$

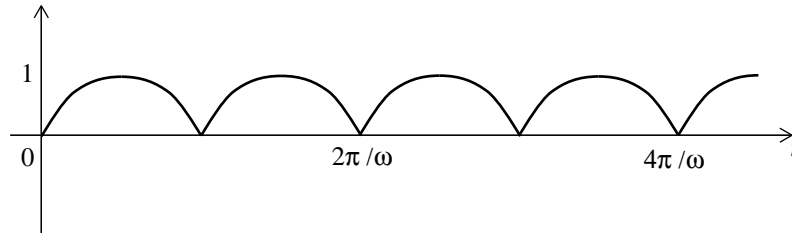


FIGURE 5.15. Rectification de l'onde de l'exemple 5.27.

Par (5.38)

$$\mathcal{L}(f)(s) = \frac{1}{1 - e^{-2\pi s/\omega}} \int_0^{\pi/\omega} e^{-st} \sin \omega t \, dt.$$

On remarque que cette intégrale est la partie imaginaire de l'intégrale

$$\begin{aligned} \int_0^{\pi/\omega} e^{(-s+i\omega)t} \, dt &= \frac{1}{-s+i\omega} e^{(-s+i\omega)t} \Big|_0^{\pi/\omega} \\ &= \frac{-s-i\omega}{s^2+\omega^2} \left(-e^{-s\pi/\omega} - 1 \right). \end{aligned}$$

Alors,

$$\mathcal{L}(f)(s) = \frac{\omega (1 + e^{-\pi s/\omega})}{(s^2 + \omega^2) (1 - e^{-2\pi s/\omega})}$$

et, par la formule

$$1 - e^{-2\pi s/\omega} = (1 + e^{-\pi s/\omega}) (1 - e^{-\pi s/\omega}),$$

on obtient

$$\mathcal{L}(f)(s) = \frac{\omega}{(s^2 + \omega^2) (1 - e^{-\pi s/\omega})}.$$

(b) Résolution par Matlab symbolique.—

```
syms pi s t omega
G = int(exp(-s*t)*sin(omega*t), t, 0, pi/omega)
G = omega*(exp(-pi/omega*s)+1)/(s^2+omega^2)
F = 1/(1-exp(-2*pi*s/omega))*G
F = 1/(1-exp(-2*pi/omega*s))*omega*(exp(-pi/omega*s)+1)/(s^2+omega^2)
```

□

EXEMPLE 5.27. Trouver la transformée de Laplace de la rectification totale de l'onde

$$f(t) = \sin \omega t$$

(V. figure 5.15).

RÉSOLUTION. L'onde rectifiée de période $p = 2\pi/\omega$ est

$$f(t) = |\sin \omega t| = \begin{cases} \sin \omega t, & \text{si } 0 < t < \pi\omega, \\ -\sin \omega t, & \text{si } \pi < t < 2\pi\omega. \end{cases}$$

Par la méthode utilisée à l'exemple 5.26, on obtient

$$\mathcal{L}(f)(s) = \frac{\omega}{s^2 + \omega^2} \coth \frac{\pi s}{2\omega}. \quad \square$$

CHAPITRE 6

Solutions séries

6.1. La méthode

On illustre la méthode de solution en série par un exemple très simple.

EXEMPLE 6.1. Soit

$$y'' + 25y = 0, \quad y(0) = 3, \quad y'(0) = 13.$$

Trouver la solution en série de puissances :

$$y(x) = a_0 + a_1x + a_2x^2 + \dots$$

RÉSOLUTION. Dans ce cas simple, on connaît déjà l'unique solution :

$$y(x) = a \cos 5x + b \sin 5x,$$

où les constantes a et b sont déterminées par les conditions initiales :

$$y(0) = a = 3 \implies a = 3,$$

$$y'(0) = 5b = 13 \implies b = \frac{13}{5}.$$

On sait aussi que

$$\cos 5x = 1 - \frac{(5x)^2}{2!} + \frac{(5x)^4}{4!} - \frac{(5x)^6}{6!} + \dots$$

et

$$\sin 5x = 5x - \frac{(5x)^3}{3!} + \frac{(5x)^5}{5!} - \frac{(5x)^7}{7!} + \dots$$

Pour obtenir la solution en série, on pose

$$y(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

dans

$$y'' + 25y = 0.$$

Alors

$$y''(x) = 2a_2 + 3 \times 2a_3x + 4 \times 3a_4x^2 + 5 \times 4a_5x^3 + 6 \times 5a_6x^4 + \dots$$

$$25y(x) = 25a_0 + 25a_1x + 25a_2x^2 + 25a_3x^3 + 25a_4x^4 + \dots$$

et l'on somme :

$$0 = (2a_2 + 25a_0) + (3 \times 2a_3 + 25a_1)x + (4 \times 3a_4 + 25a_2)x^2 + \dots, \quad \text{pour tout } x.$$

Puisque nous avons une identité en

$$1, x, x^2, x^3, \dots,$$

il suit que tous les coefficients sont nuls. Donc avec a_0 et a_1 indéterminés, on a

$$\begin{aligned} 2a_2 + 25a_0 = 0 &\implies a_2 = -\frac{5^2}{2!}a_0, \\ 3 \times 2a_3 + 25a_1 = 0 &\implies a_3 = -\frac{5^2}{3!}a_1, \\ 4 \times 3a_4 + 25a_2 = 0 &\implies a_4 = \frac{5^4}{4!}a_0, \\ 5 \times 4a_5 + 25a_3 = 0 &\implies a_5 = \frac{5^4}{5!}a_1, \end{aligned}$$

etc., d'où l'on obtient le développement

$$\begin{aligned} y(x) &= a_0 \left[1 - \frac{1}{2!}(5x)^2 + \frac{1}{4!}(5x)^4 - \frac{1}{6!}(5x)^6 + \dots \right] \\ &\quad + \frac{a_1}{5} \left[5x - \frac{1}{3!}(5x)^3 + \frac{1}{5!}(5x)^5 - \dots \right] \\ &= a_0 \cos 5x + \frac{a_1}{5} \sin 5x. \end{aligned}$$

La condition initiale $y(0) = 3$ détermine a_0 :

$$a_0 = 3,$$

et la condition initiale $y'(0) = 13$ détermine a_1 :

$$5 \frac{a_1}{5} = 13, \quad a_1 = 13. \quad \square$$

6.2. Fondements de la méthode des séries de puissances

Il sera avantageux de considérer les séries entières dans le plan complexe. On rappelle qu'un point z du plan complexe \mathbb{C} admet les représentations suivantes :

- *cartésienne* ou *algébrique* :

$$z = x + iy, \quad i^2 = -1,$$

- *trigonométrique* :

$$z = r(\cos \theta + i \sin \theta),$$

- *polaire* ou *eulérienne* :

$$z = r e^{i\theta},$$

où

$$r = \sqrt{x^2 + y^2}, \quad \theta = \arg z = \arctan \frac{y}{x}.$$

On note $\bar{z} = x - iy$ le conjugué complexe de z et

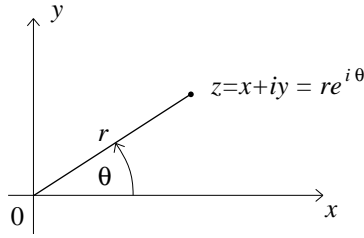
$$|z| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}} = r$$

le module de z (V. figure 6.1).

EXEMPLE 6.2. Prolonger la fonction

$$f(x) = \frac{1}{1-x},$$

au plan complexe et la développer en série de puissances de centre $z_0 = 0$, $z_0 = -1$ et $z_0 = i$.

FIGURE 6.1. Un point $z = x + iy = r e^{i\theta}$ du plan complexe \mathbb{C} .

RÉSOLUTION. On prolonge la fonction d'une variable réelle

$$f(x) = \frac{1}{1-x}, \quad x \in \mathbb{R} \setminus \{1\},$$

au plan complexe,

$$f(z) = \frac{1}{1-z}, \quad z = x + iy \in \mathbb{C}.$$

C'est une fonction rationnelle qui admet un pôle simple en $z = 1$. On dit que $z = 1$ est un pôle de $f(z)$ puisque $|f(z)|$ tend vers $+\infty$ quand $z \rightarrow 1$. De plus, il s'agit d'un pôle simple puisque $1 - z$ apparaît à la première puissance au dénominateur.

On développe $f(z)$ en série de TAYLOR près de 0 :

$$f(z) = f(0) + \frac{1}{1!} f'(0)z + \frac{1}{2!} f''(0)z^2 + \dots$$

Puisque

$$\begin{aligned} f(z) &= \frac{1}{(1-z)} && \implies f(0) = 1, \\ f'(z) &= \frac{1!}{(1-z)^2} && \implies f'(0) = 1!, \\ f''(z) &= \frac{2!}{(1-z)^3} && \implies f''(0) = 2!, \\ &\vdots && \\ f^{(n)}(z) &= \frac{n!}{(1-z)^{n+1}} && \implies f^{(n)}(0) = n!, \end{aligned}$$

il suit que

$$f(z) = \frac{1}{1-z} = 1 + z + z^2 + z^3 + \dots = \sum_{n=0}^{\infty} z^n.$$

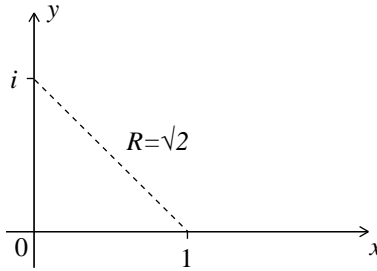
La série converge **absolument** pour $|z| \equiv \sqrt{x^2 + y^2} < 1$, c'est-à-dire

$$\sum_{n=0}^{\infty} |z|^n < \infty, \quad \text{pour tout } |z| < 1,$$

et **uniformément** pour $|z| \leq \rho < 1$, c'est-à-dire

$$\left| \sum_{n=N}^{\infty} z^n \right| < \epsilon, \quad \text{pour tout } N > N_\epsilon, \quad \text{et pour tout } |z| \leq \rho < 1.$$

Le rayon de convergence R de $\sum_{n=0}^{\infty} z^n$ est 1.

FIGURE 6.2. Distance du centre $a = i$ au pôle $z = 1$.

Maintenant, on développe f au voisinage de $z = -1$:

$$\begin{aligned} f(z) &= \frac{1}{1-z} = \frac{1}{1-(z+1-1)} \\ &= \frac{1}{2-(z+1)} = \frac{1}{2} \frac{1}{1-\frac{z+1}{2}} \\ &= \frac{1}{2} \left\{ 1 + \frac{z+1}{2} + \left(\frac{z+1}{2}\right)^2 + \left(\frac{z+1}{2}\right)^3 + \left(\frac{z+1}{2}\right)^4 + \dots \right\}. \end{aligned}$$

La série converge absolument pour

$$\left| \frac{z+1}{2} \right| < 1, \quad |z+1| < 2, \quad |z-(-1)| < 2.$$

Le centre du disque de convergence est $z = -1$ et son rayon est $R = 2$.

Enfin, on développe f au voisinage de $z = i$:

$$\begin{aligned} f(z) &= \frac{1}{1-z} = \frac{1}{1-(z-i+i)} \\ &= \frac{1}{(1-i)-(z-i)} = \frac{1}{1-i} \frac{1}{1-\frac{z-i}{1-i}} \\ &= \frac{1}{1-i} \left\{ 1 + \frac{z-i}{1-i} + \left(\frac{z-i}{1-i}\right)^2 + \left(\frac{z-i}{1-i}\right)^3 + \left(\frac{z-i}{1-i}\right)^4 + \dots \right\}. \end{aligned}$$

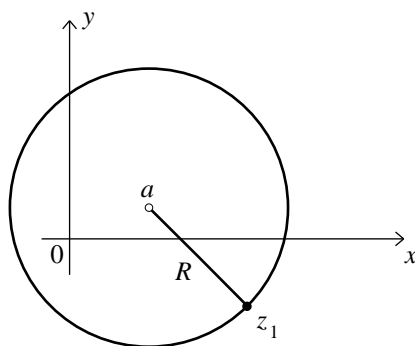
La série converge absolument pour

$$\left| \frac{z-i}{1-i} \right| < 1, \quad \text{c'est-à-dire } |z-i| < |1-i| = \sqrt{2}.$$

On voit que le centre du disque de convergence est bien $z = i$ et son rayon est $R = \sqrt{2}$ (V. figure 6.2). \square

Cet exemple montre que le développement de TAYLOR d'une fonction $f(z)$, en $z = a$ et de rayon de convergence R , n'est plus convergent dès que $|z-a| \geq R$, c'est-à-dire dès que $|z-a|$ est plus grand que la distance de a à la singularité z_1 la plus proche (V. figure 6.3).

On emploiera le résultat suivant.

FIGURE 6.3. Distance R du centre a à la singularité z_1 la plus proche.

THÉORÈME 6.1 (Critères de convergence). *L'inverse du rayon de convergence R d'une série entière de centre a ,*

$$\sum_{m=0}^{\infty} a_m (z - a)^m, \quad (6.1)$$

est égal à la limite supérieure :

$$\frac{1}{R} = \limsup_{m \rightarrow \infty} |a_m|^{1/m}. \quad (6.2)$$

On a aussi le critère :

$$\frac{1}{R} = \lim_{m \rightarrow \infty} \left| \frac{a_{m+1}}{a_m} \right|, \quad (6.3)$$

si cette limite existe.

DÉMONSTRATION. Le critère de convergence de CAUCHY affirme que la série

$$\sum_{m=0}^{\infty} c_m$$

converge si

$$\lim_{m \rightarrow \infty} |c_m|^{1/m} < 1.$$

Alors la série entière converge si

$$\lim_{m \rightarrow \infty} |a_m (z - a)^m|^{1/m} = \lim_{m \rightarrow \infty} |a_m|^{1/m} |z - a| < 1.$$

Soit R le maximum de $|z - a|$ tel que l'égalité

$$\lim_{m \rightarrow \infty} |a_m|^{1/m} R = 1$$

soit satisfaite. S'il y a plusieurs limites, il faut prendre la limite supérieure. Ceci démontre le critère (6.2). Le second critère découle du critère de D'ALEMBERT qui affirme que la série

$$\sum_{m=0}^{\infty} c_m$$

converge si

$$\lim_{m \rightarrow \infty} \frac{|c_{m+1}|}{|c_m|} < 1.$$

Donc, la série de puissance converge si

$$\lim_{m \rightarrow \infty} \frac{|a_{m+1}(z-a)^{m+1}|}{|a_m(z-a)^m|} = \lim_{m \rightarrow \infty} \frac{|a_{m+1}|}{|a_m|} |z-a| < 1.$$

Soit R le maximum de $|z-a|$ tel que l'égalité

$$\lim_{m \rightarrow \infty} \frac{|a_{m+1}|}{|a_m|} R = 1$$

soit satisfaite. Ceci démontre le critère (6.3). □

EXEMPLE 6.3. Trouver le rayon de convergence de la série

$$\sum_{m=0}^{\infty} \frac{1}{k^m} x^{3m}$$

et de sa 1ère dérivée terme à terme.

RÉSOLUTION. Par le critère (6.2),

$$\frac{1}{R} = \limsup_{m \rightarrow \infty} |a_m|^{1/m} = \lim_{m \rightarrow \infty} \left| \frac{1}{k^m} \right|^{1/3m} = \frac{1}{|k|^{1/3}}.$$

Donc le rayon de convergence de la série est

$$R = |k|^{1/3}.$$

Pour employer le critère (6.3), on pose

$$w = x^3$$

dans la série qui devient

$$\sum_{m=0}^{\infty} \frac{1}{k^m} w^m.$$

Alors le rayon de convergence, R_1 , de la nouvelle série est donné par la formule

$$\frac{1}{R_1} = \lim_{m \rightarrow \infty} \left| \frac{k^m}{k^{m+1}} \right| = \left| \frac{1}{k} \right|,$$

c'est-à-dire $R_1 = |k|$. La série originale converge pour

$$|x^3| = |w| < |k|, \quad \text{c'est-à-dire} \quad |x| < |k|^{1/3}.$$

Le rayon de convergence R' de la série dérivée,

$$\sum_{m=0}^{\infty} \frac{3m}{k^m} x^{3m-1},$$

s'obtient de la même manière :

$$\begin{aligned} \frac{1}{R'} &= \lim_{m \rightarrow \infty} \left| \frac{3m}{k^m} \right|^{1/(3m-1)} \\ &= \lim_{m \rightarrow \infty} |3m|^{1/(3m-1)} \lim_{m \rightarrow \infty} \left| \frac{1}{k^m} \right|^{(1/m)(m/(3m-1))} \\ &= \lim_{m \rightarrow \infty} \left(\frac{1}{|k|} \right)^{1/(3-1/m)} \\ &= \frac{1}{|k|^{1/3}}, \end{aligned}$$

puisque

$$\lim_{m \rightarrow \infty} |3m|^{1/(3m-1)} = 1. \quad \square$$

On voit par récurrence que toutes les séries dérivées terme à terme d'une série donnée admettent la même rayon de convergence R .

DÉFINITION 6.1. On dit que la fonction f est *analytique* dans le disque $D(a, R)$, de centre a et de rayon $R > 0$, si elle admet un développement de centre a ,

$$f(z) = \sum_{n=0}^{\infty} a_n (z - a)^n,$$

uniformément convergent dans tout disque fermé strictement contenu dans $D(a, R)$.

Le théorème suivant est une conséquence immédiate de la définition précédente.

THÉORÈME 6.2. *Une fonction f analytique dans $D(a, R)$ admet la représentation*

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (z - a)^n$$

uniformément et absolument convergente dans $D(a, R)$. De plus $f(z)$ est indéfiniment dérivable :

$$f^{(k)}(z) = \sum_{n=k}^{\infty} \frac{f^{(n)}(a)}{(n-k)!} (z - a)^{n-k}, \quad k = 0, 1, 2, \dots,$$

dans $D(a, R)$.

DÉMONSTRATION. Puisque le rayon de convergence de la série dérivée terme à terme est R , le résultat découle du fait que la série dérivée converge uniformément dans tout disque fermé strictement contenu dans $D(a, R)$ et que $f(z)$ est différentiable sur $D(a, R)$. \square

On a le théorème général suivant pour les équations différentielles linéaires à coefficients analytiques.

THÉORÈME 6.3 (Existence de solutions en série). *Soit l'équation différentielle*

$$y'' + f(x)y' + g(x)y = r(x),$$

où f , g et r sont des fonctions analytiques au voisinage de a . Si R est le minimum des rayons de convergence des développements en série entière, de centre a , de f , g et r , alors l'équation différentielle admet une solution analytique de centre a et de rayon de convergence R .

DÉMONSTRATION. La démonstration se fait par la méthode des majorantes dans le plan complexe \mathbb{C} , qui consiste à trouver une série (à coefficients positifs) absolument convergente dans $D(a, R)$,

$$\sum_{n=0}^{\infty} b_n (x - a)^n,$$

dont les coefficients bornent en module ceux de la solution

$$y(x) = \sum_{n=0}^{\infty} a_n (x - a)^n,$$

c'est-à-dire

$$|a_n| \leq b_n. \quad \square$$

On emploiera les théorèmes 6.2 et 6.3 pour obtenir les solutions d'équations différentielles en série de puissances. Dans les trois sections suivantes, on obtiendra la solution analytique de l'équation de LEGENDRE et l'on démontrera les relations d'orthogonalité des polynômes $P_n(x)$ de LEGENDRE. À la dernière section, on obtiendra les formules de quadrature gaussiennes à deux et à trois points.

EXEMPLE 6.4. Résoudre

$$y' - xy - 1 = 0, \quad y(0) = 1$$

en série de puissances.

RÉSOLUTION. Posons

$$y(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

dans l'équation différentielle :

$$\left. \begin{array}{l} y' \\ -xy \\ -1 \end{array} \right\} = \left\{ \begin{array}{l} 1a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \dots \\ -a_0x - a_1x^2 - a_2x^3 - \dots \\ -1 \end{array} \right.$$

La somme du 1er membre est nulle parce qu'on suppose que $y(x)$ est une solution de l'équation différentielle. La somme du second membre est

$$0 = (a_1 - 1) + (2a_2 - a_0)x + (3a_3 - a_1)x^2 + (4a_4 - a_2)x^3 + \dots$$

Puisque nous avons une identité en x , le coefficient de chaque terme en x^s est nul pour $s = 0, 1, 2, \dots$. De plus, puisque l'équation différentielle est du 1er ordre, un des coefficients sera indéterminé. Donc,

$$\begin{aligned} a_1 - 1 = 0 &\implies a_1 = 1, \\ 2a_2 - a_0 = 0 &\implies a_2 = \frac{a_0}{2}, \quad a_0 \text{ arbitraire,} \\ 3a_3 - a_1 = 0 &\implies a_3 = \frac{a_1}{3} = \frac{1}{3}, \\ 4a_4 - a_2 = 0 &\implies a_4 = \frac{a_2}{4} = \frac{a_0}{8}, \\ 5a_5 - a_3 = 0 &\implies a_5 = \frac{a_3}{5} = \frac{1}{15}, \end{aligned}$$

ainsi de suite. La condition initiale $y(0) = 1$ implique que $a_0 = 1$. Donc la solution est

$$y(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{8} + \frac{x^5}{15} + \dots$$

Celle-ci coïncide avec les solutions des exemples 1.19 and 1.20. □

6.3. Équation et polynômes de Legendre

On cherche la solution générale de l'équation de LEGENDRE :

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0, \quad -1 < x < 1, \quad (6.4)$$

sous forme de série de puissances de centre $a = 0$. On récrit l'équation sous forme standard $y'' + f(x)y' + g(x)y = r(x)$:

$$y'' - \frac{2x}{1-x^2}y' + \frac{n(n+1)}{1-x^2}y = 0.$$

Les coefficients,

$$f(x) = \frac{2x}{(x-1)(x+1)}, \quad g(x) = -\frac{n(n+1)}{(x-1)(x+1)},$$

admettent des pôles simples en $x = \pm 1$. Alors, ils admettent les développements en série de puissance de centre $a = 0$ et de rayon de convergence $R = 1$:

$$\begin{aligned} f(x) &= -\frac{2x}{1-x^2} = -2x[1 + x^2 + x^4 + x^6 + \dots], \quad -1 < x < 1, \\ g(x) &= \frac{n(n+1)}{1-x^2} = n(n+1)[1 + x^2 + x^4 + x^6 + \dots], \quad -1 < x < 1, \\ r(x) &= 0, \quad -\infty < x < \infty. \end{aligned}$$

On voit donc que f et g sont analytiques sur $-1 < x < 1$ et r est analytique partout.

Par le théorème 6.3, l'équation (6.4) admet deux solutions indépendantes et analytiques sur $-1 < x < 1$.

Posons

$$y(x) = \sum_{m=0}^{\infty} a_m x^m \tag{6.5}$$

dans (6.4), avec $k = n(n+1)$:

$$\left. \begin{array}{l} y'' \\ -x^2 y'' \\ -2xy' \\ ky \end{array} \right\} = \left\{ \begin{array}{l} 2 \times 1a_2 + 3 \times 2a_3x + 4 \times 3a_4x^2 + 5 \times 4a_5x^3 + \dots \\ \qquad \qquad \qquad -2 \times 1a_2x^2 - 3 \times 2a_3x^3 - \dots \\ \qquad \qquad \qquad - \quad 2a_1x - 2 \times 2a_2x^2 - 2 \times 3a_3x^3 - \dots \\ ka_0 + \quad ka_1x + \quad ka_2x^2 + \quad ka_3x^3 + \dots \end{array} \right.$$

La somme de chacun des membres est nulle puisqu'on suppose que (6.5) est une solution :

$$\begin{aligned} 0 &= (2 \times 1a_2 + ka_0) + (3 \times 2a_3 - 2a_1 + ka_1)x \\ &\quad + (4 \times 3a_4 - 2 \times 1a_2 - 2 \times 2a_2 + ka_2)x^2 \\ &\quad + \dots \\ &\quad + [(s+2)(s+1)a_{s+2} - s(s-1)a_s - 2sa_s + ka_s]x^s \\ &\quad + \dots, \quad \text{pour tout } x. \end{aligned}$$

Puisque nous avons une identité en x , chacun des coefficients de x^s , $s = 0, 1, 2, \dots$, est nul, et puisque l'équation (6.4) est du second ordre, deux des a_m seront indéterminés. On a donc

$$2!a_2 + ka_0 = 0 \implies a_2 = -\frac{n(n+1)}{2!}a_0, \quad a_0 \text{ indéterminé,}$$

$$(3 \times 2)a_3 + (-2 + k)a_1 = 0 \implies a_3 = \frac{2 - n(n+1)}{3!}a_1, \quad a_1 \text{ indéterminé,}$$

$$(s+2)(s+1)a_{s+2} + [-s(s-1) - 2s + n(n+1)]a_s = 0$$

$$\implies a_{s+2} = -\frac{(n-s)(n+s+1)}{(s+2)(s+1)}a_s, \quad s = 0, 1, 2, \dots,$$

d'où

$$a_2 = -\frac{n(n+1)}{2!}a_0, \quad a_3 = -\frac{(n-1)(n+2)}{3!}a_1, \quad (6.6)$$

$$a_4 = \frac{(n-2)n(n+1)(n+3)}{4!}a_0, \quad a_5 = \frac{(n-3)(n-1)(n+2)(n+4)}{5!}a_1, \quad (6.7)$$

etc. On peut donc écrire la solution de la forme :

$$y(x) = a_0y_1(x) + a_1y_2(x), \quad (6.8)$$

où

$$y_1(x) = 1 - \frac{n(n+1)}{2!}x^2 + \frac{(n-2)n(n+1)(n+3)}{4!}x^4 - + \dots,$$

$$y_2(x) = x - \frac{(n-1)(n+2)}{3!}x^3 + \frac{(n-3)(n-1)(n+2)(n+4)}{5!}x^5 - + \dots$$

Ces séries convergent pour $|x| < R = 1$. On remarque que y_1 est paire et y_2 est impaire. Puisque

$$\frac{y_1(x)}{y_2(x)} \neq \text{constante},$$

il suit que y_1 et y_2 sont deux solutions indépendantes et (6.8) est la solution générale.

COROLLAIRE 6.1. *Pour n pair, $y_1(x)$ est un polynôme pair,*

$$y_1(x) = k_n P_n(x),$$

et de même, pour n impair, on a le polynôme impair

$$y_2(x) = k_n P_n(x),$$

où $P_n(x)$ est le polynôme de LEGENDRE de degré n , tel que $P_n(1) = 1$.

Voici les 6 premiers polynômes de LEGENDRE (V. figure 6.4) :

$$P_0(x) = 1, \quad P_1(x) = x,$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1), \quad P_3(x) = \frac{1}{2}(5x^3 - 3x),$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3), \quad P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x).$$

On remarque que les n zéros du polynôme $P_n(x)$, de degré n , sont tous dans l'intervalle $]-1, 1[$. De plus, ils sont simples et entrelacent les $n-1$ zéros de $P_{n-1}(x)$; les zéros de fonctions orthogonales jouissent ordinairement de ces deux propriétés.

REMARQUE 6.1. On peut montrer que les séries donnant y_1 , resp. y_2 , divergent en $x = \pm 1$ si $n \neq 0, 2, 4, \dots$, resp. $n \neq 1, 3, 5, \dots$

On peut construire les polynômes de Legendre au moyen de Matlab symbolique avec $P_n(1) = 1$:

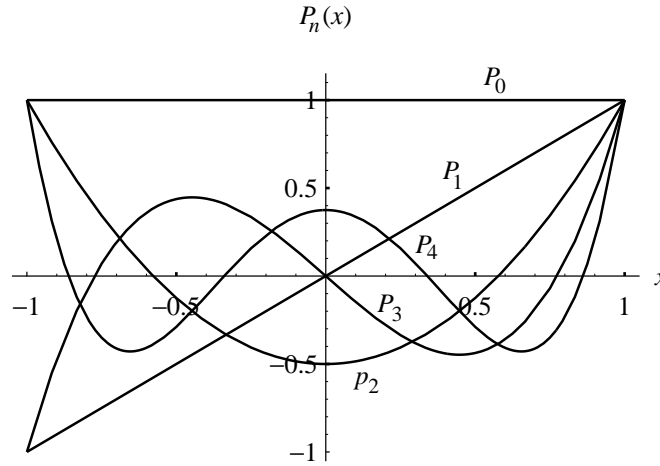


FIGURE 6.4. Les 5 premiers polynômes de LEGENDRE.

```
>> dsolve('(1-x^2)*D2y-2*x*Dy=0', 'y(1)=1', 'x')
y = 1
```

```
>> dsolve('(1-x^2)*D2y-2*x*Dy+2*y=0', 'y(1)=1', 'x')
y = x
```

```
>> dsolve('(1-x^2)*D2y-2*x*Dy+6*y=0', 'y(1)=1', 'x')
y = -1/2+3/2*x^2
```

ainsi de suite. Avec la boîte symbolique étendue de Matlab, on peut obtenir les polynômes de Legendre $P_n(x)$ en recourant au noyau complet de Maple par la commande `orthopoly[P](n,x)`, elle-même référencée par la commande `mhhelp orthopoly[P]`.

6.4. Orthogonalité des polynômes de Legendre

THÉORÈME 6.4. *Les polynômes de LEGENDRE $P_n(x)$ satisfont la relation d'orthogonalité suivante :*

$$\int_{-1}^1 P_m(x)P_n(x) dx = \begin{cases} 0, & m \neq n, \\ \frac{2}{2n+1}, & m = n. \end{cases} \quad (6.9)$$

DÉMONSTRATION. On donnera deux démonstrations de la 2^e partie. La 1^{re} partie (c'est-à-dire pour $m \neq n$) découle de l'équation de LEGENDRE :

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0,$$

réécrite sous forme de divergence :

$$L_n y := [(1-x^2)y']' + n(n+1)y = 0.$$

Puisque P_m et P_n sont solutions respectivement de $L_m y = 0$ et $L_n y = 0$, on a

$$P_n(x)L_m(P_m) = 0, \quad P_m(x)L_n(P_n) = 0.$$

On intègre ces deux expressions de -1 à 1 :

$$\int_{-1}^1 P_n(x)[(1-x^2)P'_m(x)]' dx + m(m+1) \int_{-1}^1 P_n(x)P_m(x) dx = 0,$$

$$\int_{-1}^1 P_m(x)[(1-x^2)P'_n(x)]' dx + n(n+1) \int_{-1}^1 P_m(x)P_n(x) dx = 0,$$

et l'on intègre le 1^{er} terme de chacune de ces expressions par parties :

$$P_n(x)(1-x^2)P'_m(x) \Big|_{-1}^1 - \int_{-1}^1 P'_n(x)(1-x^2)P'_m(x) dx$$

$$+ m(m+1) \int_{-1}^1 P_n(x)P_m(x) dx = 0,$$

$$P_m(x)(1-x^2)P'_n(x) \Big|_{-1}^1 - \int_{-1}^1 P'_m(x)(1-x^2)P'_n(x) dx$$

$$+ n(n+1) \int_{-1}^1 P_m(x)P_n(x) dx = 0.$$

Les deux termes intégrés sont nuls et le terme suivant de chacune des équations est identique. Donc, en soustrayant ces équations on obtient l'orthogonalité des P_n :

$$[m(m+1) - n(n+1)] \int_{-1}^1 P_m(x)P_n(x) dx = 0$$

$$\implies \int_{-1}^1 P_m(x)P_n(x) dx = 0 \quad \text{pour } m \neq n.$$

La 2^e partie, pour $m = n$, suit de la **formule de RODRIGUES** :

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n]. \quad (6.10)$$

En effet,

$$\int_{-1}^1 P_n^2(x) dx = \frac{1}{2^n} \times \frac{1}{n!} \times \frac{1}{2^n} \times \frac{1}{n!} \int_{-1}^1 \left[\frac{d^n}{dx^n} (x^2 - 1)^n \right] \left[\frac{d^n}{dx^n} (x^2 - 1)^n \right] dx$$

{et intégrant par parties n fois:}

$$= \frac{1}{2^n} \times \frac{1}{n!} \times \frac{1}{2^n} \times \frac{1}{n!} \left[\frac{d^{n-1}}{dx^{n-1}} (x^2 - 1)^n \frac{d^n}{dx^n} (x^2 - 1) \right] \Big|_{-1}^1$$

$$+ (-1)^1 \int_{-1}^1 \frac{d^{n-1}}{dx^{n-1}} (x^2 - 1)^n \frac{d^{n+1}}{dx^{n+1}} (x^2 - 1)^n dx$$

+ ...

$$= \frac{1}{2^n} \times \frac{1}{n!} \times \frac{1}{2^n} \times \frac{1}{n!} (-1)^n \int_{-1}^1 (x^2 - 1)^n \frac{d^{2n}}{dx^{2n}} (x^2 - 1)^n dx$$

{et dérivant $2n$ fois:}

$$= \frac{1}{2^n} \times \frac{1}{n!} \times \frac{1}{2^n} \times \frac{1}{n!} (-1)^n (2n)! \int_{-1}^1 1 \times (x^2 - 1)^n dx$$

{et intégrant de nouveau par parties n fois:}

$$\begin{aligned}
&= \frac{(-1)^n (2n)!}{2^n n! 2^n n!} \left[\frac{x}{1} (x^2 - 1)^n \Big|_{-1}^1 + \frac{(-1)^1}{1!} 2n \int_{-1}^1 x^2 (x^2 - 1)^{n-1} dx \right] \\
&\quad + \dots \\
&= \frac{(-1)^n (2n)!}{2^n n! 2^n n!} (-1)^n \frac{2n 2(n-1) 2(n-2) \dots 2(n-(n-1))}{1 \times 3 \times 5 \times \dots \times (2n-1)} \int_{-1}^1 x^{2n} dx \\
&= \frac{(-1)^n (-1)^n (2n)!}{2^n n! 2^n n!} \frac{2^n n!}{1 \times 3 \times 5 \times \dots \times (2n-1)} \frac{1}{(2n+1)} x^{2n+1} \Big|_{-1}^1 \\
&= \frac{2}{2n+1}. \quad \square
\end{aligned}$$

REMARQUE 6.2. La formule de RODRIGUES se démontre par un calcul direct avec $n = 0, 1, 2, 3, \dots$, ou autrement. On calcule $P_4(x)$ au moyen de la formule de Rodrigues avec la commande `diff` de Matlab symbolique.

```

>> syms x f p4
>> f = (x^2-1)^4
    f = (x^2-1)^4
>> p4 = (1/(2^4*prod(1:4)))*diff(f,x,4)
    p4 = x^4+3*(x^2-1)*x^2+3/8*(x^2-1)^2
>> p4 = expand(p4)
    p4 = 3/8-15/4*x^2+35/8*x^4

```

On présente une 2^e démonstration de la formule :

$$\|P_n\|^2 := \int_{-1}^1 [P_n(x)]^2 dx = \frac{2}{2n+1}$$

au moyen de la **fonction génératrice** de $P_n(x)$:

$$\sum_{k=0}^{\infty} P_k(x) t^k = \frac{1}{\sqrt{1-2xt+t^2}}. \quad (6.11)$$

DÉMONSTRATION. Élevons au carré chacun des deux membres :

$$\sum_{k=0}^{\infty} P_k^2(x) t^{2k} + \sum_{j \neq k} P_j(x) P_k(x) t^{j+k} = \frac{1}{1-2xt+t^2},$$

et intégrons en x sur $[-1, 1]$:

$$\sum_{k=0}^{\infty} \left[\int_{-1}^1 P_k^2(x) dx \right] t^{2k} + \sum_{j \neq k} \left[\int_{-1}^1 P_j(x) P_k(x) dx \right] t^{j+k} = \int_{-1}^1 \frac{dx}{1-2xt+t^2}.$$

Comme P_j et P_k sont orthogonaux pour $j \neq k$, le 2^e terme du 1^{er} membre est nul et nous obtenons après intégration du 2^e membre :

$$\begin{aligned}
\sum_{k=0}^{\infty} \|P_k\|^2 t^{2k} &= -\frac{1}{2t} \ln(1-2xt+t^2) \Big|_{x=-1}^{x=1} \\
&= -\frac{1}{t} [\ln(1-t) - \ln(1+t)].
\end{aligned}$$

On multiplie par t :

$$\sum_{k=0}^{\infty} \|P_k\|^2 t^{2k+1} = -\ln(1-t) + \ln(1+t)$$

et l'on dérive par rapport à t :

$$\begin{aligned} \sum_{k=0}^{\infty} (2k+1) \|P_k\|^2 t^{2k} &= \frac{1}{1-t} + \frac{1}{1+t} \\ &= \frac{2}{1-t^2} \\ &= 2(1+t^2+t^4+t^6+\dots), \quad \text{pour tout } t, |t| < 1. \end{aligned}$$

Puisque nous avons une identité en t , on peut identifier les coefficients de t^{2k} :

$$(2k+1) \|P_k\|^2 = 2 \implies \|P_k\|^2 = \frac{2}{2k+1}. \quad \square$$

REMARQUE 6.3. On peut obtenir la fonction génératrice (6.11) en développant le 2^e membre en une série de TAYLOR en puissances de t , ce qui se fait facilement par la commande `taylor` de Matlab symbolique :

```
>> syms t x; f = 1/(1-2*x*t+t^2)^(1/2);
>> g = taylor(f,3,t)
g = 1+t*x+(-1/2+3/2*x^2)*t^2
      +(-3/2*x+5/2*x^3)*t^3+(3/8-15/4*x^2+35/8*x^4)*t^4
```

6.5. Série de Fourier–Legendre

Soit le développement de $f(x)$, $-1 < x < 1$, en série de FOURIER–LEGENDE :

$$f(x) = \sum_{m=0}^{\infty} a_m P_m(x), \quad -1 < x < 1,$$

où

$$a_m = \frac{2m+1}{2} \int_{-1}^1 f(x) P_m(x) dx.$$

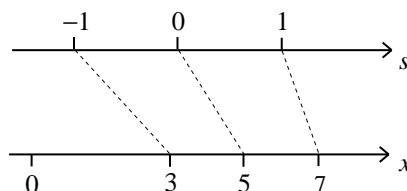
La valeur des a_m suit des relations d'orthogonalité (6.9) des $P_m(x)$ sur $-1 < x < 1$:

$$\begin{aligned} \int_{-1}^1 f(x) P_n(x) dx &= \sum_{m=0}^{\infty} a_m \int_{-1}^1 P_m(x) P_n(x) dx \\ &= a_m \int_{-1}^1 P_m(x) P_m(x) dx = \frac{2}{2m+1} a_m. \end{aligned}$$

Pour les polynômes $p(x)$ de degré k , on obtient un développement fini,

$$f(x) = \sum_{m=0}^k a_m P_m(x)$$

sur $-\infty < x < \infty$, sans intégration, par un simple changement de base : de x^m en $P_m(x)$ pour $m = 1, 2, \dots, k$.

FIGURE 6.5. Application affine de $x \in [3, 7]$ sur $s \in [-1, 1]$.

EXEMPLE 6.5. Développer le polynôme :

$$p(x) = x^3 - 2x^2 + 4x + 1$$

sur $[-1, 1]$ selon les polynômes de LEGENDRE $P_0(x), P_1(x), \dots$

RÉSOLUTION. On exprime les puissances de x suivant la base des polynômes de LEGENDRE

$$P_0(x) = 1 \implies 1 = P_0(x),$$

$$P_1(x) = x \implies x = P_1(x),$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1) \implies x^2 = \frac{2}{3}P_2(x) + \frac{1}{3}P_0(x),$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x) \implies x^3 = \frac{2}{5}P_3(x) + \frac{3}{5}P_1(x).$$

On évite ainsi le calcul d'intégrales. Alors

$$\begin{aligned} p(x) &= \frac{2}{5}P_3(x) + \frac{3}{5}P_1(x) - \frac{4}{3}P_2(x) - \frac{2}{3}P_0(x) + 4P_1(x) + P_0(x) \\ &= \frac{2}{5}P_3(x) - \frac{4}{3}P_2(x) + \frac{23}{5}P_1(x) + \frac{1}{3}P_0(x). \quad \square \end{aligned}$$

EXEMPLE 6.6. Développer le polynôme

$$p(x) = 2 + 3x + 5x^2$$

sur $[3, 7]$ selon les polynômes de LEGENDRE $P_0(x), P_1(x), \dots$

RÉSOLUTION. Appliquons $x \in [3, 7]$ sur $s \in [-1, 1]$ (V. figure 6.5). Posons $s = ax + b$. Alors

$$-1 = 3a + b \quad \text{et} \quad 1 = 7a + b \implies a = \frac{1}{2} \quad \text{et} \quad b = -\frac{5}{2}.$$

On obtient les applications affines suivantes réciproques l'une de l'autre :

$$s = \frac{x - 5}{2} \quad \text{et} \quad x = 2s + 5. \quad (6.12)$$

Alors

$$\begin{aligned} p(x) &= p(2s + 5) \\ &= 2 + 3(2s + 5) + 5(2s + 5)^2 \\ &= 142 + 106s + 20s^2 \\ &= 142P_0(s) + 106P_1(s) + 20 \left[\frac{2}{3}P_2(s) + \frac{1}{3}P_0(s) \right]. \end{aligned}$$

Donc

$$p(x) = \left(142 + \frac{20}{3}\right) P_0\left(\frac{x-5}{2}\right) + 106P_1\left(\frac{x-5}{2}\right) + \frac{40}{3}P_2\left(\frac{x-5}{2}\right). \quad \square$$

EXEMPLE 6.7. Calculer les trois premiers termes du développement de FOURIER–LEGENDRE de la fonction

$$f(x) = \begin{cases} 0, & -1 < x < 0, \\ x, & 0 < x < 1. \end{cases}$$

RÉSOLUTION. Posons

$$f(x) = \sum_{m=0}^{\infty} a_m P_m(x), \quad -1 < x < 1.$$

Alors,

$$a_m = \frac{2m+1}{2} \int_{-1}^1 f(x) P_m(x) dx.$$

Donc,

$$\begin{aligned} a_0 &= \frac{1}{2} \int_{-1}^1 f(x) P_0(x) dx = \frac{1}{2} \int_0^1 x dx = \frac{1}{4}, \\ a_1 &= \frac{3}{2} \int_{-1}^1 f(x) P_1(x) dx = \frac{3}{2} \int_0^1 x^2 dx = \frac{1}{2}, \\ a_2 &= \frac{5}{2} \int_{-1}^1 f(x) P_2(x) dx = \frac{5}{2} \int_0^1 x \frac{1}{2} (3x^2 - 1) dx = \frac{5}{16}. \end{aligned}$$

On a donc l'approximation

$$f(x) \approx \frac{1}{4} P_0(x) + \frac{1}{2} P_1(x) + \frac{5}{16} P_2(x). \quad \square$$

EXEMPLE 6.8. Calculer les trois premiers termes du développement de FOURIER–LEGENDRE de la fonction

$$f(x) = e^x, \quad 0 \leq x \leq 1.$$

RÉSOLUTION. Pour utiliser l'orthogonalité des polynômes de LEGENDRE on transforme le domaine de $f(x)$ de $[0, 1]$ à $[-1, 1]$ au moyen de la substitution

$$s = 2\left(x - \frac{1}{2}\right), \quad \text{c'est-à-dire } x = \frac{s}{2} + \frac{1}{2}.$$

Alors,

$$f(x) = e^x = e^{(1+s)/2} = \sum_{m=0}^{\infty} a_m P_m(s), \quad -1 \leq s \leq 1,$$

où

$$a_m = \frac{2m+1}{2} \int_{-1}^1 e^{(1+s)/2} P_m(s) ds.$$

On calcule d'abord trois intégrales par récurrence :

$$\begin{aligned}
 I_0 &= \int_{-1}^1 e^{s/2} ds = 2 \left(e^{1/2} - e^{-1/2} \right), \\
 I_1 &= \int_{-1}^1 s e^{s/2} ds = 2s e^{s/2} \Big|_{-1}^1 - 2 \int_{-1}^1 e^{s/2} ds \\
 &= 2 \left(e^{1/2} + e^{-1/2} \right) - 2I_0 \\
 &= -2e^{1/2} + 6e^{-1/2}, \\
 I_2 &= \int_{-1}^1 s^2 e^{s/2} ds = 2s^2 e^{s/2} \Big|_{-1}^1 - 4 \int_{-1}^1 s e^{s/2} ds \\
 &= 2 \left(e^{1/2} - e^{-1/2} \right) - 4I_1 \\
 &= 10e^{1/2} - 26e^{-1/2}.
 \end{aligned}$$

Il suit que

$$\begin{aligned}
 a_0 &= \frac{1}{2} e^{1/2} I_0 = e - 1 \approx 1.7183, \\
 a_1 &= \frac{3}{2} e^{1/2} I_1 = -3e + 9 \approx 0.8452, \\
 a_2 &= \frac{5}{2} e^{1/2} \frac{1}{2} (3I_2 - I_0) = 35e - 95 \approx 0.1399.
 \end{aligned}$$

On a donc l'approximation

$$f(x) \approx 1.7183P_0(2x-1) + 0.8452P_1(2x-1) + 0.1399P_2(2x-1). \quad \square$$

6.6. Une application : la quadrature gaussienne

On obtient facilement les formules de quadrature gaussienne à n points au moyen des polynômes de LEGENDRE. Par souci de simplicité, on ne considère que $n = 2$ et $n = 3$. On précise que le nombre n de points se réfère aux n points en lesquels on évalue la fonction à intégrer sur l'intervalle standardisé $[-1, 1]$, et non pas au nombre de sous-intervalles en lesquels on a l'habitude de diviser l'intervalle d'intégration $[a, b]$ afin de diminuer l'erreur de la valeur numérique de l'intégrale.

EXEMPLE 6.9. Déterminer les 4 paramètres de la quadrature gaussienne à 2 points :

$$\int_{-1}^1 f(x) dx = af(x_1) + bf(x_2).$$

RÉSOLUTION. Par symétrie, on prévoit que les nœuds sont opposés, $x_1 = -x_2$, et les poids sont égaux, $a = b$. Puisqu'on a 4 paramètres, la formule est exacte pour les polynômes de degré 3 et, par l'exemple 6.5, il suffit de considérer les polynômes $P_0(x), \dots, P_3(x)$. Comme $P_0(x) = 1$ est orthogonal à $P_n(x)$, $n =$

1, 2, ..., on a

$$2 = \int_{-1}^1 P_0(x) dx = aP_0(x_1) + bP_0(x_2) = a + b, \quad (6.13)$$

$$0 = \int_{-1}^1 1 \times P_1(x) dx = aP_1(x_1) + bP_1(x_2) = ax_1 + bx_2, \quad (6.14)$$

$$0 = \int_{-1}^1 1 \times P_2(x) dx = aP_2(x_1) + bP_2(x_2), \quad (6.15)$$

$$0 = \int_{-1}^1 1 \times P_3(x) dx = aP_3(x_1) + bP_3(x_2), \quad (6.16)$$

Pour satisfaire (6.15) on choisit x_1 et x_2 tels que

$$P_2(x_1) = P_2(x_2) = 0,$$

c'est-à-dire :

$$P_2(x) = \frac{1}{2}(3x^2 - 1) = 0 \Rightarrow -x_1 = x_2 = \frac{1}{\sqrt{3}} = 0.577\ 350\ 27.$$

Alors, par (6.14)

$$a = b.$$

De plus, (6.16) est automatiquement satisfaite car $P_3(x)$ est impair. Enfin, par (6.13),

$$a = b = 1.$$

On a donc la formule de GAUSS à deux points :

$$\int_{-1}^1 f(x) dx = f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right). \quad \square \quad (6.17)$$

EXEMPLE 6.10. Déterminer les 6 paramètres de la quadrature gaussienne à 3 points :

$$\int_{-1}^1 f(x) dx = af(x_1) + bf(x_2) + cf(x_3).$$

RÉSOLUTION. Par symétrie, on prévoit que les nœuds extrêmes sont opposés, $x_1 = -x_3$, et $x_2 = 0$, les poids extrêmes sont égaux, $a = c$, et le poids central dépasse les deux autres, $b > a = c$. Puisqu'on a 6 paramètres, la formule est exacte pour les polynômes de degré 5 et, par l'exemple 6.5, il suffit de considérer

la base $P_0(x), \dots, P_5(x)$:

$$2 = \int_{-1}^1 P_0(x) dx = aP_0(x_1) + bP_0(x_2) + cP_0(x_3), \quad (6.18)$$

$$0 = \int_{-1}^1 P_1(x) dx = aP_1(x_1) + bP_1(x_2) + cP_1(x_3), \quad (6.19)$$

$$0 = \int_{-1}^1 P_2(x) dx = aP_2(x_1) + bP_2(x_2) + cP_2(x_3), \quad (6.20)$$

$$0 = \int_{-1}^1 P_3(x) dx = aP_3(x_1) + bP_3(x_2) + cP_3(x_3), \quad (6.21)$$

$$0 = \int_{-1}^1 P_4(x) dx = aP_4(x_1) + bP_4(x_2) + cP_4(x_3), \quad (6.22)$$

$$0 = \int_{-1}^1 P_5(x) dx = aP_5(x_1) + bP_5(x_2) + cP_5(x_3). \quad (6.23)$$

Pour satisfaire (6.21), prenons pour x_1, x_2, x_3 les 3 zéros de

$$P_3(x) = \frac{1}{2}(5x^3 - 3x) = \frac{1}{2}x(5x^2 - 3),$$

c'est-à-dire

$$-x_1 = x_3 = \sqrt{\frac{3}{5}} = 0.774\,596\,7, \quad x_2 = 0.$$

Alors (6.19) implique

$$-\sqrt{\frac{3}{5}}a + \sqrt{\frac{3}{5}}c = 0 \Rightarrow a = c;$$

donc (6.23) est satisfaite puisque $P_5(x)$ est impair. De plus, par la substitution $a = c$ dans (6.20), on obtient

$$a \frac{1}{2} \left(3 \times \frac{3}{5} - 1 \right) + b \left(-\frac{1}{2} \right) + a \frac{1}{2} \left(3 \times \frac{3}{5} - 1 \right) = 0,$$

c'est-à-dire

$$4a - 5b + 4a = 0 \quad \text{ou} \quad 8a - 5b = 0. \quad (6.24)$$

Maintenant, de (6.18) on déduit

$$2a + b = 2 \quad \text{ou} \quad 10a + 5b = 10. \quad (6.25)$$

Si l'on additionne la seconde expression de (6.24) et de (6.25), on obtient

$$a = \frac{10}{18} = \frac{5}{9} = 0.555;$$

alors

$$b = 2 - \frac{10}{9} = \frac{8}{9} = 0.888.$$

Enfin, on vérifie que (6.22) est satisfaite. Puisque

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3),$$

on a

$$\begin{aligned} 2 \times \frac{5 \times 1}{9 \times 8} \left(35 \times \frac{9}{25} - 30 \times \frac{3}{5} + 3 \right) + \frac{8}{9} \times \frac{3}{8} &= \frac{2 \times 5}{9 \times 8} \left(\frac{315 - 450 + 75}{25} \right) + \frac{8}{9} \times \frac{3}{8} \\ &= \frac{2 \times 5}{9 \times 8} \times \frac{(-60)}{25} + \frac{8 \times 3}{9 \times 8} \\ &= \frac{-24 + 24}{9 \times 8} = 0. \end{aligned}$$

On a donc la formule de GAUSS à trois points :

$$\int_{-1}^1 f(x) dx = \frac{5}{9} f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9} f(0) + \frac{5}{9} f\left(\sqrt{\frac{3}{5}}\right). \quad \square \quad (6.26)$$

REMARQUE 6.4. On a normalisé l'intervalle d'intégration des quadratures de GAUSS sur $[-1, 1]$. Pour intégrer sur $[a, b]$ on emploie le changement de variable déjà utilisée à l'exemple 6.8 :

$$x = \frac{(b-a)t + b + a}{2}, \quad dx = \left(\frac{b-a}{2}\right) dt.$$

Alors

$$\int_a^b f(x) dx = \frac{b-a}{2} \int_{-1}^1 f\left(\frac{(b-a)t + b + a}{2}\right) dt.$$

EXEMPLE 6.11. Évaluer

$$I = \int_0^{\pi/2} \sin x dx$$

par la formule de GAUSS à 2 point appliquée une seule fois sur tout l'intervalle $[0, \pi/2]$ et sur les demi-intervalles $[0, \pi/4]$ et $[\pi/4, \pi/2]$.

RÉSOLUTION. Posons

$$x = \frac{(\pi/2)t + \pi/2}{2}, \quad dx = \frac{\pi}{4} dt.$$

En $t = -1$, $x = 0$ et, en $t = 1$, $x = \pi/2$. Alors

$$\begin{aligned} I &= \frac{\pi}{4} \int_{-1}^1 \sin\left(\frac{\pi t + \pi}{4}\right) dt \\ &\approx \frac{\pi}{4} [1.0 \times \sin(0.10566\pi) + 1.0 \times \sin(0.39434\pi)] \\ &= 0.99847. \end{aligned}$$

L'erreur est 1.53×10^{-3} . Sur les demi-intervalles, on a

$$\begin{aligned} I &= \frac{\pi}{8} \int_{-1}^1 \sin\left(\frac{\pi t + \pi}{8}\right) dt + \frac{\pi}{8} \int_{-1}^1 \sin\left(\frac{\pi t + 3\pi}{8}\right) dt \\ &\approx \frac{\pi}{8} \left[\sin\frac{\pi}{8} \left(-\frac{1}{\sqrt{3}} + 1\right) + \sin\frac{\pi}{8} \left(\frac{1}{\sqrt{3}} + 1\right) \right. \\ &\quad \left. + \sin\frac{\pi}{8} \left(-\frac{1}{\sqrt{3}} + 3\right) + \sin\frac{\pi}{8} \left(\frac{1}{\sqrt{3}} + 3\right) \right] \\ &= 0.99991016676989. \end{aligned}$$

L'erreur est 8.983×10^{-5} . Voici une solution obtenue par Matlab. Par soucis de généralité, définissons la fonction fichier M `exp5_10.m`,

```
function f=exp5_10(t)
% evaluate the function f(t)
f=sin(t);
```

Voici le programme de la quadrature gaussienne à deux points :

```
>> clear
>> a = 0; b = pi/2; c = (b-a)/2; d= (a+b)/2;
>> weight = [1 1]; node = [-1/sqrt(3) 1/sqrt(3)];
>> syms x t
>> x = c*node+d;
>> nv1 = c*weight*exp5_10(x)' % valeur numerique de l'integrale
nv1 = 0.9985
>> error1 = 1 - nv1 % l'erreur dans la solution
error1 = 0.0015
```

On obtient l'autre partie de la même façon. □

On évalue l'intégrale de l'exemple 6.11 par la méthode de Simpson adaptée (`quad`) et la méthode de Newton-Cotes à 8 panneaux adaptée (`quad8`) de Matlab.

```
>> v1 = quad('sin',0,pi/2)
v1 = 1.00000829552397
>> v2 = quad8('sin',0,pi/2)
v2 = 1.00000000000000
```

à une erreur relative au 10^{-3} près.

REMARQUE 6.5. La quadrature gaussienne est la formule la plus précise pour un nombre de points donné. L'erreur pour la formule à n points est

$$E_n(f) = \frac{2}{(2n+1)!} \left[\frac{2^n (n!)^2}{(2n)!} \right]^2 f^{(2n)}(\xi), \quad -1 < \xi < 1.$$

La formule à n points est donc exacte pour un polynôme de degré $2n - 1$.

Partie 2

Méthodes numériques

Résolution d'équations nonlinéaires par récurrence

7.1. Calculs arithmétiques sur ordinateurs

7.1.1. Définitions. On utilisera la terminologie suivante.

- (1) Si a est la valeur exacte d'un calcul et \tilde{a} est une valeur approchée pour le même calcul, alors

$$\epsilon = \tilde{a} - a$$

est l'**erreur** en \tilde{a} et $|\epsilon|$ est l'**erreur absolue**. Si $a \neq 0$,

$$\epsilon_r = \frac{\tilde{a} - a}{a} = \frac{\epsilon}{a}$$

est l'**erreur relative** en \tilde{a} .

- (2) Les **bornes supérieures** pour les erreurs absolue et relative en \tilde{a} sont les nombre B_a et B_r tels que

$$|\epsilon| = |\tilde{a} - a| < B_a, \quad |\epsilon_r| = \left| \frac{\tilde{a} - a}{a} \right| < B_r.$$

- (3) On appelle **erreur d'arrondi** ou **erreur de chute** l'erreur causée par l'arrondi ou la troncature d'un nombre réel après un nombre fini de chiffres à droite de la virgule décimale, en particulier sur ordinateur (V. Subsection 7.1.2).
- (4) Dans les calculs scientifiques en **virgule flottante**, on représente un nombre c de longueur d dans la base β par l'expression

$$c = \pm 0.b_1 b_2 \cdots b_d \times \beta^N,$$

où $b_1 \neq 0$, $0 \leq b_i < \beta$. On appelle $b_1 b_2 \cdots b_d$ la **mantisse** ou **partie décimale** et N l'**exposant** of c . Par exemple, si $d = 5$ et $\beta = 10$, on a

$$0.27120 \times 10^2, \quad -0.31224 \times 10^3.$$

- (5) Le nombre de **chiffres significatifs** d'un nombre en virgule flottante est le nombre de chiffres du premier au dernier chiffres non nuls. Par exemple, si $d = 4$ et $\beta = 10$, le nombre de chiffres significatifs des trois nombres:

$$0.1203 \times 10^2, \quad 0.1230 \times 10^{-2}, \quad 0.1000 \times 10^3,$$

est 4, 3 et 1.

- (6) On appelle **erreur de méthode** l'erreur causée par troncature d'une série infinie après un nombre fini de termes.

REMARQUE 7.1. Pour simplifier l'écriture on écrira souvent les nombres en virgule flottante avec des zéros immédiatement à droite de la virgule décimale ou avec des chiffres non nuls à gauche de cette virgule et sans exposants :

$$0.001203, \quad 12300.04$$

7.1.2. Nombres arrondis et nombres tronqués. On arrondit les nombres réels en direction opposée de l'origine. Par exemple, en base 10, on arrondit

$$c = \pm 0.b_1b_2 \dots b_d \times 10^N$$

à k chiffres de la façon suivante:

(i) Si $0.b_{k+1}b_{k+2} \dots b_m \geq 0.5$, on arrondit c à

$$(0.b_1b_2 \dots b_{k-1}b_k + 0.1 \times 10^{-k+1}) \times 10^N.$$

(ii) Si $0.b_{k+1}b_{k+2} \dots b_m < 0.5$, on arrondit c à

$$0.b_1b_2 \dots b_{k-1}b_k \times 10^N.$$

EXEMPLE 7.1. Nombres arrondis à 3 chiffres :

$$1.9234542 \approx 1.92$$

$$2.5952100 \approx 2.60$$

$$1.9950000 \approx 2.00$$

$$-4.9850000 \approx -4.99$$

On tronque à k chiffres un nombre en virgule flottante en remplaçant les chiffres à droite du k ème chiffre par des zéros.

7.1.3. Annulation dans les calculs. La soustraction de deux chiffres presque égaux conduit à une perte du nombre de chiffres significatifs. Il est mieux d'éviter cette perte que d'essayer d'estimer l'erreur qui en résulte. L'exemple 7.2 illustre ce problème.

EXEMPLE 7.2. Résoudre l'équation du second degré

$$x^2 - 1634x + 2 = 0$$

avec arrondi à 10 décimales.

RÉSOLUTION. On a

$$x = 817 \pm \sqrt{2\,669\,948},$$

par la formule ordinaire. Donc,

$$x_1 = 817 + 816.998\,776\,0 = 1.633\,998\,776 \times 10^3,$$

$$x_2 = 817 - 816.998\,776\,0 = 1.224\,000\,000 \times 10^{-3}.$$

Les 4 premiers zéros des 6 zéros de la mantisse de x_2 proviennent d'annulation et sont sans valeur. On obtient un résultat plus précis pour x_2 par la relation

$$x_1x_2 = 2.$$

Dans ce cas,

$$x_2 = 1.223\,991\,125 \times 10^{-3},$$

où tous les chiffres sont significatifs. □

On trouve à l'exemple 7.2 une formule numériquement stable pour résoudre l'équation du second degré

$$ax^2 + bx + c = 0, \quad a \neq 0,$$

c'est-à-dire

$$x_1 = \frac{1}{2a} \left[-b - \text{sign}(b) \sqrt{b^2 - 4ac} \right], \quad x_2 = \frac{c}{ax_1},$$

où sign dénote la fonction signum :

$$\text{sign}(x) = \begin{cases} +1, & \text{if } x \geq 0, \\ -1, & \text{if } x < 0. \end{cases}$$

EXEMPLE 7.3. Soit 4.81 la valeur de x et 12.752 la valeur de y arrondies à 3 chiffres. Trouver le plus petit intervalle qui contient la valeur exacte de $x - y$.

RÉSOLUTION. Puisque

$$4.805 \leq x < 4.815 \quad \text{et} \quad 12.7515 \leq y < 12.7525,$$

alors

$$4.805 - 12.7525 \leq x - y < 4.815 - 12.7515 \Leftrightarrow -7.9475 \leq x - y < -7.9365. \quad \square$$

EXEMPLE 7.4. Trouver l'erreur et l'erreur relative dans les approximations rationnelles bien connues $22/7$ et $355/113$ du nombre transcendant π et exprimer la réponse en virgule flottante à 3 chiffres.

RÉSOLUTION. L'erreur et l'erreur relative dans $22/7$ sont

$$\epsilon = 22/7 - \pi, \quad \epsilon_r = \epsilon/\pi,$$

qu'on calcule avec Matlab :

```
pp = pi
pp = 3.14159265358979
r1 = 22/7.
r1 = 3.14285714285714
abserr1 = r1 - pi
abserr1 = 0.00126448926735
relerr1 = abserr1/pi
relerr1 = 4.024994347707008e-04
```

Donc, l'erreur et l'erreur relative dans $22/7$ arrondies à 3 chiffres sont respectivement

$$\epsilon = 0.126 \times 10^{-2} \quad \text{et} \quad \epsilon_r = 0.402 \times 10^{-3}.$$

De nouveau par Matlab l'erreur et l'erreur relative dans $355/113$ sont

```
r2 = 355/113.
r2 = 3.14159292035398
abserr2 = r2 - pi
abserr2 = 2.667641894049666e-07
relerr2 = abserr2/pi
relerr2 = 8.491367876740610e-08
```

Donc, l'erreur et l'erreur relative dans 355/113 arrondies à 3 chiffres sont respectivement

$$\epsilon = 0.267 \times 10^{-6} \quad \text{et} \quad \epsilon_r = 0.849 \times 10^{-7}. \quad \square$$

7.2. Quelques théorème du calcul différentiel et intégral

Les résultats suivants sont nécessaires pour justifier les méthodes que l'on va introduire dans la suite.

THÉORÈME 7.1 (Théorème des valeurs intermédiaires). *Soit $a < b$ et f une fonction continue sur l'intervalle $[a, b]$. Si d est un nombre strictement entre $f(a)$ et $f(b)$, alors il existe c tel que $a < c < b$ et $f(c) = d$.*

COROLLAIRE 7.1. *Soit $a < b$ et f une fonction continue sur l'intervalle $[a, b]$. Si $f(a)f(b) < 0$ alors il existe un zéro de f dans l'intervalle $]a, b[$.*

DÉMONSTRATION. Comme $f(a)$ et $f(b)$ sont de signes différents, 0 est entre $f(a)$ et $f(b)$. On utilise le théorème des valeurs intermédiaires avec $w = 0$. \square

THÉORÈME 7.2 (Théorème des valeurs extrêmes). *Soit $a < b$ et f une fonction continue sur l'intervalle $[a, b]$. Ils existent $\alpha \in [a, b]$ et $\beta \in [a, b]$ telles que pour tout $x \in [a, b]$*

$$f(\alpha) \leq f(x) \leq f(\beta).$$

THÉORÈME 7.3 (Théorème des accroissements finis). *Soit $a < b$ et f une fonction continue sur l'intervalle $[a, b]$ et dérivable sur l'intervalle $]a, b[$. Il existe c tel que $a < c < b$ et*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

THÉORÈME 7.4 (Théorème des valeurs moyennes pour intégrales). *Soit $a < b$ et f une fonction continue sur l'intervalle $[a, b]$. Si g est une fonction intégrable sur $[a, b]$ et g ne change pas de signe sur l'intervalle $[a, b]$, alors il existe c tel que $a < c < b$ et*

$$\int_a^b f(x)g(x) dx = f(c) \int_a^b g(x) dx.$$

On a un théorème semblable pour les sommes.

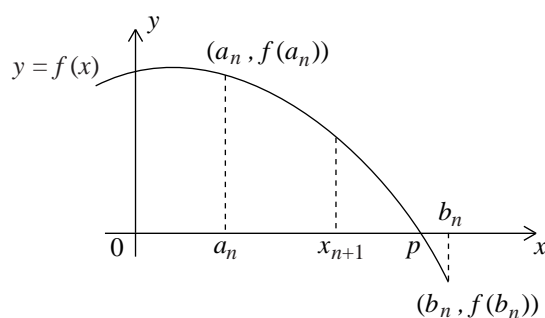
THÉORÈME 7.5 (Théorème des valeurs moyennes pour les sommes). *Soit $\{w_i\}$, $i = 1, 2, \dots, n$, un ensemble de n nombres réels distincts et f une fonction continue sur l'intervalle $[a, b]$. Si les nombres w_i sont tous du même signe et tous les points $x_i \in [a, b]$, alors il existe un nombre $c \in [a, b]$ tel que*

$$\sum_{i=1}^n w_i f(x_i) = f(c) \sum_{i=1}^n w_i.$$

7.3. Méthode de la dichotomie

La méthode de la dichotomie, appelée aussi méthode de bisection, construit une suite d'intervalles de longueur décroissante qui contiennent une racine p de $f(x) = 0$. Si

$$f(a)f(b) < 0 \quad \text{et} \quad f \quad \text{est continue} \quad [a, b],$$

FIGURE 7.1. Le n -ième pas de la méthode de la dichotomie.

alors, par le corollaire 7.1, $f(x) = 0$ admet une racine entre a et b . La racine est soit entre

$$a \text{ et } \frac{a+b}{2}, \text{ si } f(a)f\left(\frac{a+b}{2}\right) < 0,$$

soit entre

$$\frac{a+b}{2} \text{ et } b, \text{ si } f\left(\frac{a+b}{2}\right)f(b) < 0,$$

ou exactement en

$$\frac{a+b}{2}, \text{ si } f\left(\frac{a+b}{2}\right) = 0.$$

La figure 7.1 illustre le n ième pas de la méthode de la dichotomie.

L'algorithme de la **méthode de la dichotomie** :

ALGORITHME 7.1 (Méthode de la dichotomie). Soit $f(x)$ continue sur $[a, b]$ et $f(a)f(b) < 0$.

- (1) Choisir $a_0 = a$, $b_0 = b$; la tolérance TOL ; le nombre maximum d'itérations N_0 .
- (2) Pour $n = 0, 1, 2, \dots, N_0$, calculer

$$x_{n+1} = \frac{a_n + b_n}{2}.$$
- (3) Si $f(x_{n+1}) = 0$ ou $(b_n - a_n)/2 < TOL$, alors sortir $p (= x_{n+1})$ et arrêter.
- (4) Sinon $f(x_{n+1})$ et $f(a_n)$ sont de signes opposés; poser $a_{n+1} = a_n$ et $b_{n+1} = x_{n+1}$.
- (5) Sinon poser $a_{n+1} = x_{n+1}$ et $b_{n+1} = b_n$.
- (6) Répéter (2), (3), (4) et (5).
- (7) Sortie 'La méthode a flanché après N_0 itérations'; arrêter.

On trouve d'autres critères d'arrêt à la sous-section 7.4.1. Le taux de convergence de la dichotomie est faible mais la méthode converge toujours.

La fonction M en Matlab tirée de <ftp://ftp.cs.cornell.edu/pub/cv> fait la dichotomie.

```
function root = Bisection(fname,a,b,delta)
%
% Pre:
% fname string that names a continuous function f(x) of
```

```

%          a single variable.
%
%      a,b  define an interval [a,b]
%          f is continuous, f(a)f(b) < 0
%
%      delta non-negative real number.
%
% Post:
%      root  the midpoint of an interval [alpha,beta]
%          with the property that f(alpha)f(beta)<=0 and
%          |beta-alpha| <= delta+eps*max(|alpha|,|beta|)
%
fa = feval(fname,a);
fb = feval(fname,b);
if fa*fb > 0
    disp('Initial interval is not bracketing.')
    return
end
if nargin==3
    delta = 0;
end
while abs(a-b) > delta+eps*max(abs(a),abs(b))
    mid = (a+b)/2;
    fmid = feval(fname,mid);
    if fa*fmid<=0
        % There is a root in [a,mid].
        b = mid;
        fb = fmid;
    else
        % There is a root in [mid,b].
        a = mid;
        fa = fmid;
    end
end
root = (a+b)/2;

```

EXEMPLE 7.5. Approcher $\sqrt{2}$ par dichotomie. Arrêter quand $|x_{n+1} - x_n| < 10^{-2}$.

RÉSOLUTION. On cherche une racine de $f(x) = x^2 - 2 = 0$. On démarre avec $a_0 = 1$ et $b_0 = 2$ et l'on obtient par dichotomie

$$x_{n+1} = \frac{a_n + b_n}{2}.$$

Les résultats se trouvent dans le tableau 7.1. La réponse est $\sqrt{2} \approx 1.414063$ à 10^{-2} près. On remarque que la racine est dans l'intervalle $[1.414063, 1.421875]$. \square

EXEMPLE 7.6. Montrer que la fonction $f(x) = x^3 + 4x^2 - 10$ admet un zéro unique dans l'intervalle $[1, 2]$, faire 8 itérations de dichotomie sur f et borner l'erreur absolue.

TABLE 7.1. Résultats de l'exemple 7.5.

| n | x_n | a_n | b_n | $ x_{n-1} - x_n $ | $f(x_n)$ | $f(a_n)$ |
|-----|----------|----------|----------|-------------------|----------|----------|
| 0 | | 1 | 2 | | | — |
| 1 | 1.500000 | 1 | 1.500000 | .500000 | + | — |
| 2 | 1.250000 | 1.250000 | 1.500000 | .250000 | — | — |
| 3 | 1.375000 | 1.375000 | 1.500000 | .125000 | — | — |
| 4 | 1.437500 | 1.375000 | 1.437500 | .062500 | + | — |
| 5 | 1.406250 | 1.406250 | 1.437500 | .031250 | — | — |
| 6 | 1.421875 | 1.406250 | 1.421875 | .015625 | + | — |
| 7 | 1.414063 | 1.414063 | 1.421875 | .007812 | — | — |

TABLE 7.2. Résultats de l'exemple 7.6.

| n | x_n | a_n | b_n | $f(x_n)$ | $f(a_n)$ |
|-----|-------------|-------------|-------------|----------|----------|
| 0 | | 1 | 2 | | — |
| 1 | 1.500000000 | 1 | 1.500000000 | + | — |
| 2 | 1.250000000 | 1.250000000 | 1.500000000 | — | — |
| 3 | 1.375000000 | 1.250000000 | 1.375000000 | + | — |
| 4 | 1.312500000 | 1.312500000 | 1.375000000 | — | — |
| 5 | 1.343750000 | 1.343750000 | 1.375000000 | — | — |
| 6 | 1.359375000 | 1.359375000 | 1.375000000 | — | — |
| 7 | 1.367187500 | 1.359375000 | 1.367187500 | + | — |
| 8 | 1.363281250 | 1.363281250 | 1.367187500 | — | — |

RÉSOLUTION. Puisque

$$f(1) = -5 < 0 \quad \text{et} \quad f(2) = 14 > 0,$$

alors $f(x)$ admet un zéro, p , dans $[1, 2]$. Ce zéro est unique puisque $f(x)$ est strictement croissante sur $[1, 2]$. De fait

$$f'(x) = 3x^2 + 4x > 0 \quad \text{pour tout } x \text{ entre 1 et 2.}$$

Les résultats se trouvent dans le tableau 7.2.

Après 8 itérations, on voit que p est entre 1.363 281 250 et 1.367 187 500. Donc l'erreur absolue en p est bornée par

$$1.367 187 500 - 1.363 281 250 = 0.003 906 25. \quad \square$$

EXEMPLE 7.7. Trouver le nombre d'itérations minimum pour avoir à l'exemple 7.6 une erreur absolue inférieure à 10^{-4} .

RÉSOLUTION. Puisque le zéro, p , est dans chaque intervalle $[a_n, b_n]$, après n itérations l'erreur est au plus $b_n - a_n$. On cherche donc n tel que $b_n - a_n < 10^{-4}$. Puisque chaque itération coupe l'intervalle en deux, on voit facilement que

$$b_n - a_n = (2 - 1)/2^n.$$

Donc n doit satisfaire l'inégalité

$$2^{-n} < 10^{-4},$$

c'est-à-dire,

$$\ln 2^{-n} < \ln 10^{-4}, \quad \text{ou} \quad -n \ln 2 < -4 \ln 10.$$

Alors,

$$n > 4 \ln 10 / \ln 2 = 13.28771238 \implies n = 14.$$

Donc 14 itérations suffisent. \square

7.4. Récurrence de point fixe

Soit $f(x)$ une fonction à valeur réelle de la variable réelle x . Dans cette section, on présente des méthodes de récurrence pour résoudre des équations de la forme

$$f(x) = 0. \quad (7.1)$$

Une **racine** de l'équation $f(x) = 0$, ou un **zéro** de $f(x)$, est un nombre p tel que $f(p) = 0$.

Pour trouver une racine de (7.1), on récrit cette équation sous forme équivalente

$$x = g(x), \quad (7.2)$$

telle que $g(p) = p$ et l'on dit que p est un point fixe de g . Par exemple, $g(x) = x - f(x)$.

On dit que (7.1) et (7.2) sont **équivalentes** (sur un intervalle donné) si chaque racine de (7.1) dans cet intervalle est un point fixe de (7.2) et réciproquement.

On dit que la méthode de point fixe définie par la récurrence

$$x_{n+1} = g(x_n), \quad n = 0, 1, \dots, \quad (7.3)$$

est convergente si, pour une valeur initiale x_0 , la suite x_0, x_1, \dots converge sur le nombre p . Si $g(x)$ est continue, alors $p = g(p)$, comme on peut le voir en faisant tendre $n \rightarrow \infty$ dans (7.3). On appelle p un **point fixe** de la récurrence (7.3).

On voit facilement que les équations

$$x^3 + 9x - 9 = 0, \quad x = (9 - x^3)/9$$

sont équivalentes. Le problème, c'est de bien choisir la fonction $g(x)$ et la valeur de départ x_0 pour avoir convergence. À cet effet, on doit définir différents types de points fixes.

DÉFINITION 7.1. On dit qu'un point fixe, $p = g(p)$, de la récurrence

$$x_{n+1} = g(x_n)$$

est *attractif*, *répulsif* ou *indifférent* si le *multiplicateur* $g'(p)$ de $g(x)$ en p satisfait respectivement

$$|g'(p)| < 1, \quad |g'(p)| > 1, \quad \text{ou} \quad |g'(p)| = 1.$$

THÉORÈME 7.6 (Théorème du point fixe). *Soit $g(x)$ une fonction à valeur réelle telle que :*

- (1) $g(x) \in [a, b]$ pour tout $x \in [a, b]$;
- (2) $g(x)$ est dérivable sur $[a, b]$;
- (3) il existe un nombre K , $0 < K < 1$, tel que $|g'(x)| \leq K$ pour tout $x \in]a, b[$.

Alors $g(x)$ admet un unique point fixe $p \in [a, b]$. De plus, pour $x_0 \in [a, b]$ arbitraire, la suite x_0, x_1, x_2, \dots définie par

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, \dots,$$

converge vers p .

DÉMONSTRATION. Si $g(a) = a$ ou $g(b) = b$, l'existence d'un point fixe attractif est évidente. Dans le cas contraire, il suit que $g(a) > a$ et $g(b) < b$. On définit la fonction auxiliaire

$$h(x) = g(x) - x.$$

Alors h est continue sur $[a, b]$ et

$$h(a) = g(a) - a > 0, \quad h(b) = g(b) - b < 0.$$

Par le corollaire 7.1, il existe un nombre $p \in]a, b[$ tel que $h(p) = 0$, c'est-à-dire, $g(p) = p$. Donc p est un point fixe de $g(x)$.

Pour démontrer l'unicité on suppose que p et q sont deux points fixes distincts de $g(x)$ sur $[a, b]$. Par le théorème des accroissements finis 7.3, il existe un nombre c entre p et q (et donc dans $[a, b]$) tel que

$$|p - q| = |g(p) - g(q)| = |g'(c)| |p - q| \leq K |p - q| < |p - q|,$$

ce qui est une contradiction. Donc $p = q$ et le point fixe attractif dans $[a, b]$ est unique.

On démontre la convergence. Par le théorème des accroissements finis 7.3, pour chaque paire de nombre x et y dans $[a, b]$, il existe un nombre c entre x et y tel que

$$g(x) - g(y) = g'(c)(x - y).$$

Donc,

$$|g(x) - g(y)| \leq K|x - y|.$$

En particulier,

$$|x_{n+1} - p| = |g(x_n) - g(p)| \leq K|x_n - p|.$$

Après $n + 1$ répétitions de ce processus, on a

$$|x_{n+1} - p| \leq K^{n+1}|x_0 - p| \rightarrow 0, \quad \text{quand } n \rightarrow \infty,$$

puisque $0 < K < 1$. Donc la suite $\{x_n\}$ converge vers p . □

EXEMPLE 7.8. Trouver une racine de l'équation

$$f(x) = x^3 + 9x - 9 = 0$$

dans l'intervalle $[0, 1]$ par une récurrence de point fixe.

RÉSOLUTION. Résoudre l'équation équivaut à trouver un point fixe de

$$g(x) = (9 - x^3)/9.$$

Puisque

$$f(0)f(1) = -9 < 0,$$

par le corollaire 7.1, $f(x)$ admet un zéro, p , entre 0 et 1. La condition (3) du théorème 7.6 est satisfaite avec $K = 1/3$ puisque

$$|g'(x)| = |-x^2/3| \leq 1/3$$

pour tout x entre 0 et 1. Les autres conditions sont aussi satisfaites.

TABLE 7.3. Résultats de l'exemple 7.8.

| n | x_n | Erreur ϵ_n | $\epsilon_n/\epsilon_{n-1}$ |
|-----|-------------------|---------------------|-----------------------------|
| 0 | 0.500000000000000 | -0.41490784153366 | 1.000000000000000 |
| 1 | 0.986111111111111 | 0.07120326957745 | -0.17161225325185 |
| 2 | 0.89345451579409 | -0.02145332573957 | -0.30129691890395 |
| 3 | 0.92075445888550 | 0.00584661735184 | -0.27252731920515 |
| 4 | 0.91326607850598 | -0.00164176302768 | -0.28080562295762 |
| 5 | 0.91536510274262 | 0.00045726120896 | -0.27851839836463 |

On fait 5 itérations à partir de $x_0 = 0.5$ avec Matlab et la fonction M `exp7_8.m` :

```
function x1 = exp7_8(x0); % Exemple 7.8.
x1 = (9-x0^3)/9;
```

On obtient la solution exacte

$$0.91490784153366$$

au moyen de quelques 30 itérations par le processus suivant : résout le problème.

```
xexact = 0.91490784153366;
N = 5; x=zeros(N+1,4);
x0 = 0.5; x(1,:) = [0 x0 (x0-xexact), 1];
for i = 1:N
xt=exp7_8(x(i,2));
x(i+1,:) = [i xt (xt-xexact), (xt-xexact)/x(i,3)];
end
```

Le tableau 7.3 présente les itérées, leurs erreurs et le rapport des erreurs successives. On voit que ce rapport est presque constant. Il suit que l'ordre de convergence, défini à la sous-section 7.4.2, est égal à un. \square

On montre à l'exemple 7.9 que la convergence de la récurrence $x_{n+1} = g(x_n)$ vers un point fixe dépend du réarrangement judicieux de l'équation $f(x) = 0$ à résoudre.

Une récurrence de point fixe peut admettre des cycles en plus de points fixes. Pour définir un cycle, on utilise la notation : $g^2(x) = g(g(x))$, $g^3(x) = g(g^2(x))$ etc.

DÉFINITION 7.2. Soit la récurrence

$$x_{n+1} = g(x_n).$$

Un k -cycle de $g(x)$ est un ensemble de k points distincts,

$$x_0, x_1, x_2, \dots, x_{k-1},$$

qui satisfont les relations

$$x_1 = g(x_0), x_2 = g^2(x_0), \dots, x_{k-1} = g^{k-1}(x_0), x_0 = g^k(x_0).$$

Le *multipliateur* d'un k cycle est le nombre réel ou complexe

$$(g^k)'(x_j) = g'(x_{k-1}) \cdots g'(x_0), \quad j = 0, 1, \dots, k-1.$$

On dit qu'un k -cycle est *attractif*, *répulsif*, ou *indifférent* respectivement si

$$|(g^k)'(x_j)| < 1, \quad > 1, \quad = 1.$$

Un point fixe est un 1-cycle.

On voit que le multiplicateur d'un cycle est le même en chaque point du cycle.

EXEMPLE 7.9. Trouver une racine de l'équation

$$f(x) = x^3 + 4x^2 - 10 = 0$$

dans l'intervalle $[1, 2]$ par diverses récurrences de point fixe et étudier leur convergence.

RÉSOLUTION. Puisque $f(1)f(2) = -70 < 0$, l'équation $f(x) = 0$ admet une racine dans l'intervalle $[1, 2]$. On obtient la racine exacte par la commande `roots` de Matlab :

```
p=[1 4 0 -10]; % le polynome f(x)
r =roots(p)
r =
-2.68261500670705 + 0.35825935992404i
-2.68261500670705 - 0.35825935992404i
1.36523001341410
```

On note x_∞ la racine réelle dans l'intervalle $[1, 2]$. Les 2 autres racines sont conjuguées complexes l'une de l'autre.

Soit cinq réarrangements $x_{n+1} = g_j(x_n)$, $j = 1, 2, 3, 4, 5$, de l'équation $f(x) = 0$ et la valeur des dérivées $g'_j(x)$ en la racine réelle $x_\infty \approx 1.365$.

$$\begin{aligned} x = g_1(x) &=: 10 + x - 4x^2 - x^3, & g'_1(x_\infty) &\approx -15.51, \\ x = g_2(x) &=: \sqrt{(10/x) - 4x}, & g'_2(x_\infty) &\approx -3.42, \\ x = g_3(x) &=: \frac{1}{2}\sqrt{10 - x^3}, & g'_3(x_\infty) &\approx -0.51, \\ x = g_4(x) &=: \sqrt{10/(4+x)}, & g'_4(x_\infty) &\approx -0.13 \\ x = g_5(x) &=: x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}, & g'_5(x_\infty) &= 0. \end{aligned}$$

On itère 6 fois chacun des réarrangements par Matlab. On emploie la fonction M de Matlab `exp7_9.m`

```
function y = exp7_9(x); % Example 7.9.
y = [10+x(1)-4*x(1)^2-x(1)^3; sqrt((10/x(2))-4*x(2));
sqrt(10-x(3)^3)/2; sqrt(10/(4+x(4)));
x(5)-(x(5)^3+4*x(5)^2-10)/(3*x(5)^2+8*x(5))];
```

et le processus suivant :

```
N = 6; x=zeros(N+1,5);
x0 = 1.5; x(1,:) = [0 x0 x0 x0 x0];
for i = 1:N
xt=exp7_9(x(i,2:5));
x(i+1,:) = [i xt];
end
```

TABLE 7.4. Résultats de l'exemple 7.9.

| | $g_1(x)$ | $g_2(x)$ | $g_3(x)$ | $g_4(x)$ | $g_5(x)$ |
|-----|--------------------------|----------------------|----------------------|---------------------|---|
| n | $10 + x - 4x^2 - x^3$ | $\sqrt{(10/x) - 4x}$ | $0.5\sqrt{10 - x^3}$ | $\sqrt{10/(4 + x)}$ | $x - \frac{x^3 + 4x^2 - 10}{2x^2 + 8x}$ |
| 0 | 1.5 | 1.5 | 1.5 | 1.5 | 1.5 |
| 1 | -0.8750 | 0.816 | 1.286953 | 1.348399 | 1.373333333 |
| 2 | 6.732421875 | 2.996 | 1.402540 | 1.367376 | 1.365262015 |
| 3 | -4.6972001×10^2 | $0.00 - 2.94i$ | 1.345458 | 1.364957 | 1.365230014 |
| 4 | 1.0275×10^8 | $2.75 - 2.75i$ | 1.375170 | 1.365264 | 1.365230013 |
| 5 | -1.08×10^{24} | $1.81 - 3.53i$ | 1.360094 | 1.365225 | |
| 6 | 1.3×10^{72} | $2.38 - 3.43i$ | 1.367846 | 1.365230 | |

On résume les résultats dans le tableau 7.4. On voit dans le tableau que x_∞ est un point fixe de $g_3(x)$, $g_4(x)$ et $g_5(x)$. De plus, $g_4(x_n)$ converge plus rapidement vers la racine 1.365 230 013 que $g_3(x_n)$, et $g_5(x)$ converge encore plus vite. De fait, ces 3 récurrences produisent un résultat exact à 10 chiffres près respectivement en 30, 15 et 4 itérations, D'autre part la suite $g_2(x_n)$ est prise dans le 2-cycle attractif

$$z_\pm = 2.27475487839820 \pm 3.60881272309733i,$$

de multiplicateur

$$g_2'(z_+)g_2'(z_-) = 0.19790433047378$$

inférieur à 1 en module. Une fois prise dans un cycle attractif, une récurrence ne peut converger vers un point fixe. Enfin, x_∞ est un point fixe répulsif de $g_1(x)$ et $x_{n+1} = g_1(x_n)$ diverge vers $-\infty$. \square

REMARQUE 7.2. Une récurrence qui démarre dans le bassin d'attraction d'un point (ou cycle) attractif, converge vers ce point fixe (ou ce cycle). Une récurrence qui démarre dans le voisinage d'un point (ou cycle) répulsif, ne peut converger vers ce point (ou ce cycle). La convergence vers un point fixe indifférent est très lente, mais on peut l'accélérer par divers processus d'accélération.

7.4.1. Critères d'arrêt. Il y a 3 critères d'arrêt usuels d'une récurrence pour trouver un zéro de $f(x)$:

- (1) après N itérations (pour N donné);
- (2) quand $|x_{n+1} - x_n| < \epsilon$ (pour ϵ donné).
- (3) quand $|f(x_n)| < \eta$ (pour η donné).

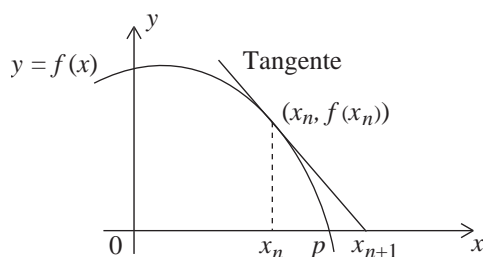
L'utilité de ces critères dépend du problème en question.

7.4.2. Ordre et taux de convergence d'une récurrence. On est souvent intéressé au taux de convergence d'une récurrence. On suppose que la fonction $g(x)$ de la récurrence

$$x_{n+1} = g(x_n)$$

admet le point fixe p ($p = g(p)$) et un développement de Taylor de centre p . On note

$$\epsilon_n = x_n - p.$$

FIGURE 7.2. Le n -ième pas de la méthode de Newton.

Alors,

$$\begin{aligned} x_{n+1} &= g(x_n) = g(p + \epsilon_n) = g(p) + g'(p)\epsilon_n + \frac{g''(p)}{2!}\epsilon_n^2 + \dots \\ &= p + g'(p)\epsilon_n + \frac{g''(p)}{2!}\epsilon_n^2 + \dots \end{aligned}$$

Donc,

$$\epsilon_{n+1} = x_{n+1} - p = g'(p)\epsilon_n + \frac{g''(p)}{2!}\epsilon_n^2 + \dots \quad (7.4)$$

DÉFINITION 7.3. L'ordre de convergence de la récurrence $x_{n+1} = g(x_n)$ est l'ordre de la première dérivée non nulle de $g(x)$ en p . Une méthode d'ordre p admet un **taux de convergence** égal à p .

À l'exemple 7.9, les récurrences $g_3(x)$ et $g_4(x)$ convergent d'ordre 1, alors que $g_5(x)$ converge d'ordre 2.

Dans le cas d'une convergence du 2ème ordre on a :

$$\frac{\epsilon_{n+1}}{\epsilon_n^2} \approx \frac{g''(p)}{2} = \text{constante.}$$

7.5. Méthodes de Newton, de la sécante et de la position fausse

7.5.1. Méthode de Newton. Soit x_n une approximation de la racine p de $f(x) = 0$. On trace la tangente

$$y = f(x_n) + f'(x_n)(x - x_n)$$

à la courbe $y = f(x)$ en $(x_n, f(x_n))$ (V. la figure 7.2). L'équation de la tangente évaluée au point de son intersection $(x_{n+1}, 0)$ avec l'axe des x ,

$$0 = f(x_n) + f'(x_n)(x_{n+1} - x_n),$$

détermine l'abscisse x_{n+1} de ce point. Si $f'(x_n) \neq 0$, la solution de cette équation pour x_{n+1} donne la **méthode de Newton**, appelée aussi **méthode de Newton-Raphson**,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \quad (7.5)$$

On remarque que la méthode de Newton est une récurrence de point fixe de la forme

$$x_{n+1} = g(x_n), \quad \text{où } g(x) = x - \frac{f(x)}{f'(x)}.$$

TABLE 7.5. Résultats de l'exemple 7.10.

| n | x_n | $ x_n - x_{n-1} $ |
|-----|----------|-------------------|
| 0 | 2 | |
| 1 | 1.5 | 0.5 |
| 2 | 1.416667 | 0.083333 |
| 3 | 1.414216 | 0.002451 |
| 4 | 1.414214 | 0.000002 |

TABLE 7.6. Résultats de l'exemple 7.11.

| n | x_n | $ x_n - x_{n-1} $ |
|-----|------------------|---------------------------|
| 0 | 1.5 | |
| 1 | 1.37333333333333 | 0.126667 |
| 2 | 1.36526201487463 | 0.00807132 |
| 3 | 1.36523001391615 | 0.000032001 |
| 4 | 1.3652300134141 | 5.0205×10^{-10} |
| 5 | 1.3652300134141 | 2.22045×10^{-16} |
| 6 | 1.3652300134141 | 2.22045×10^{-16} |

EXEMPLE 7.10. Approcher $\sqrt{2}$ par la méthode de Newton. Arrêter dès que $|x_{n+1} - x_n| < 10^{-4}$.

RÉSOLUTION. On veut résoudre l'équation

$$f(x) = x^2 - 2 = 0$$

pour $x > 0$. Dans ce cas, la méthode de Newton devient

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{x_n^2 + 2}{2x_n}.$$

Avec $x_0 = 2$, on trouve le résultat dans le tableau 7.5 avec 4 itérations. Donc,

$$\sqrt{2} \approx 1.414214.$$

On remarque que le nombre de zéros dans l'erreur double, ou à peu près, puisque le zéro est simple et donc la méthode converge d'ordre 2. \square

EXEMPLE 7.11. Approcher une racine $p \in [1, 2]$ du polynôme

$$f(x) = x^3 + 4x^2 - 10 = 0$$

de l'exemple 7.9 au moyen de 6 itérations de la méthode de Newton.

RÉSOLUTION. Dans ce cas, la méthode de Newton devient

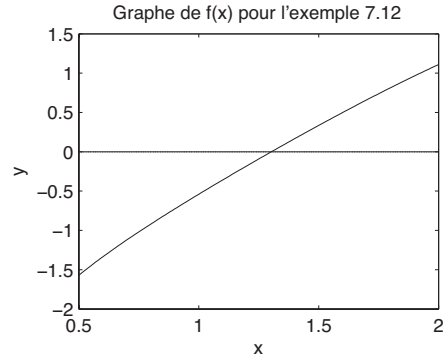
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^3 + 4x_n^2 - 10}{3x_n^2 + 8x_n} = \frac{2(x_n^3 + 2x_n^2 + 5)}{3x_n^2 + 8x_n}.$$

Avec $x_0 = 1.5$ on a les résultats dans le tableau 7.6. \square

EXEMPLE 7.12. Résoudre l'équation :

$$\ln x = \cos x$$

à 6 décimales près par Newton.

FIGURE 7.3. Graphe de $\ln x - \cos x$ pour l'exemple 7.12.

RÉSOLUTION. Soit $f(x) = \ln x - \cos x$. Alors $f(x) = 0$ quand $\ln x = \cos x$. Le graphique montre que la solution est entre $x = 1$ et $x = \pi/2$. Alors on démarre en $x_0 = 1$. On a

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{\ln x_n - \cos x_n}{(1/x_n) + \sin x_n}.$$

Donc

$$\begin{aligned} x_0 &= 1, \\ x_1 &= 1 - \frac{\ln 1 - \cos 1}{(1/1) + \sin 1} = 1.293\,408, \\ x_2 &= 1.302\,956 \\ x_3 &= 1.302\,964, \\ x_4 &= 1.302\,964, \\ &\text{stop} \end{aligned}$$

On vérifie que la réponse est correcte :

$$\begin{aligned} \ln 1.302\,964 &\approx 0.264\,641\,6, \\ \cos 1.302\,964 &\approx 0.264\,641\,6. \end{aligned}$$

Alors la solution, à 6 décimales près, est 1.302 964. \square

THÉORÈME 7.7. Soit p une racine simple de $f(x) = 0$, c'est-à-dire, $f(p) = 0$ et $f'(p) \neq 0$. Si $f''(p)$ existe, alors la méthode de Newton converge au moins d'ordre 2 avec x_0 près de p .

DÉMONSTRATION. On dérive la fonction

$$g(x) = x - \frac{f(x)}{f'(x)}$$

ce qui donne

$$\begin{aligned} g'(x) &= 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} \\ &= \frac{f(x)f''(x)}{(f'(x))^2}. \end{aligned}$$

TABLE 7.7. Résultats de l'exemple 7.13.

| n | Méthode de Newton | | Newton modifiée | |
|-----|-------------------|-----------------------------|-------------------|-------------------------------|
| | x_n | $\epsilon_{n+1}/\epsilon_n$ | x_n | $\epsilon_{n+1}/\epsilon_n^2$ |
| 0 | 0.000 | | 0.000000000000000 | |
| 1 | 0.400 | 0.600 | 0.800000000000000 | -0.2000 |
| 2 | 0.652 | 2.245 | 0.98461538461538 | -0.3846 |
| 3 | 0.806 | 0.143 | 0.99988432620012 | -0.4887 |
| 4 | 0.895 | 0.537 | 0.99999999331095 | -0.4999 |
| 5 | 0.945 | 0.522 | 1 | |
| 6 | 0.972 | 0.512 | 1 | |

Puisque $f(p) = 0$, on obtient

$$g'(p) = 0.$$

Alors, la méthode de Newton converge d'ordre 2 près d'un zéro simple de f . \square

REMARQUE 7.3. Soit la 2ème dérivée de $g(x)$:

$$g''(x) = \frac{(f'(x))^2 f''(x) + f(x) f'(x) f'''(x) - 2f(x)(f''(x))^2}{(f'(x))^3}.$$

Si $f'''(p)$ existe, on obtient

$$g''(p) = -\frac{f''(p)}{f'(p)}.$$

Donc, par (7.4), les erreurs successives satisfont l'approximation :

$$\epsilon_{n+1} \approx -\frac{1}{2} \frac{f''(p)}{f'(p)} \epsilon_n^2.$$

Ceci explique le fait que le nombre de zéros en position supérieure dans l'erreur de la méthode de Newton double près d'une racine simple de $f(x) = 0$.

EXEMPLE 7.13. Approcher la racine double $x = 1$ du polynôme

$$f(x) = (x - 1)^2(x - 2)$$

par 6 itérations des méthodes de Newton et de Newton modifiée :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad x_{n+1} = x_n - 2 \frac{f(x_n)}{f'(x_n)}.$$

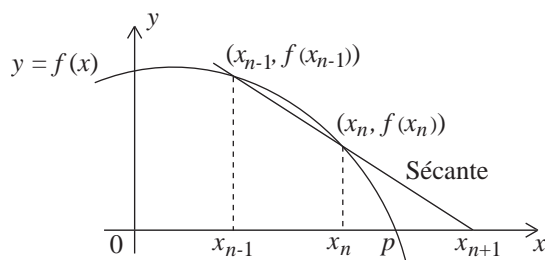
RÉSOLUTION. On a les 2 fonctions de récurrence :

$$g_1(x) = x - \frac{(x-1)(x-2)}{2(x-2) + (x-1)}, \quad g_2(x) = x - \frac{(x-1)(x-2)}{(x-2) + (x-1)}.$$

On démarre en $x_0 = 0$. On trouve les résultats dans le tableau 7.7. On voit que la méthode de Newton converge d'ordre 1 près d'un zéro double de $f(x)$ alors que la méthode de Newton modifiée converge d'ordre 2. De fait, près d'une racine de multiplicité m la méthode de Newton modifiée

$$x_{n+1} = x_n - m \frac{f(x_n)}{f'(x_n)}$$

converge d'ordre 2. \square

FIGURE 7.4. Le n -ième pas de la méthode de la sécante.

En générale la méthode de Newton peut converger vers la racine recherchée, ou vers une autre racine, ou vers un cycle attractif, surtout dans le plan complexe.

7.5.2. Méthode de la sécante. Soit x_{n-1} et x_n deux valeurs approchées de la racine p de $f(x) = 0$. On trace la sécante de la courbe $y = f(x)$ en $(x_{n-1}, f(x_{n-1}))$ et $(x_n, f(x_n))$. L'équation de la sécante est

$$y = f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} (x - x_n).$$

La $(n + 1)$ -ème itérée x_{n+1} est l'abscisse du point d'intersection $(x_{n+1}, 0)$ de la sécante avec l'axe des x (V. la figure. 7.4),

$$0 = f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} (x_{n+1} - x_n).$$

Si l'on résout pour x_{n+1} , on obtient la **méthode de la sécante**:

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n). \quad (7.6)$$

On donne l'algorithme de cette méthode :

ALGORITHME 7.2 (Méthode de la sécante). Soit la fonction $f(x)$ continue sur $[a, b]$ qui admet un zéro dans $[a, b]$.

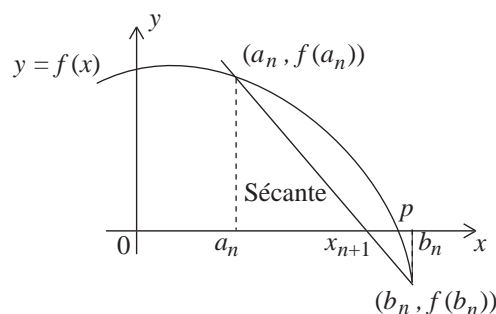
- (1) Choisir x_0 et x_1 près de la racine p .
- (2) Étant donné x_{n-1} et x_n , on obtient x_{n+1} par la formule

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n),$$

si $f(x_n) - f(x_{n-1}) \neq 0$. Sinon, on modifie les valeurs de démarrage x_0 et x_1 .

- (3) Répéter (2) jusqu'à la satisfaction du critère d'arrêt choisi (V. la sous-section 7.4.1).

La méthode de la sécante converge d'ordre 1.618 vers une racine simple. Elle peut bien ne pas converger vers une racine multiple. Cette méthode est généralement plus lente que la méthode de Newton, mais elle n'utilise pas la 1ère dérivée de $f(x)$. Dans les applications de la méthode de Newton on approche souvent $f'(x)$ numériquement par la pente de la sécante.

FIGURE 7.5. Le n -ième pas de la méthode de la position fausse.

7.5.3. Méthode de la position fausse. La méthode de la position fausse, appelée aussi *regula falsi*, ressemble à la méthode de la sécante, mais avec la condition additionnelle que pour chaque $n = 0, 1, 2, \dots$, la paire de valeurs approchées, a_n et b_n , de la racine p de $f(x) = 0$ doit satisfaire l'inégalité $f(a_n)f(b_n) < 0$. On détermine l'itérée suivante, x_{n+1} , par l'intersection de la sécante en les points $(a_n, f(a_n))$ et $(b_n, f(b_n))$ avec l'axe des x .

L'équation de la sécante en $(a_n, f(a_n))$ et $(b_n, f(b_n))$ (V. la figure Fig. 7.5) est

$$y = f(a_n) + \frac{f(b_n) - f(a_n)}{b_n - a_n} (x - a_n).$$

Donc x_{n+1} satisfait l'équation

$$0 = f(a_n) + \frac{f(b_n) - f(a_n)}{b_n - a_n} (x_{n+1} - a_n),$$

ce qui donne la **méthode de la position fausse**:

$$x_{n+1} = \frac{a_n f(b_n) - b_n f(a_n)}{f(b_n) - f(a_n)}. \quad (7.7)$$

L'algorithme de la méthode de la position fausse :

ALGORITHME 7.3 (Méthode de la position fausse). Soit $f(x)$ continue sur $[a, b]$ et $f(a)f(b) < 0$.

- (1) Choisir $a_0 = a$ et $b_0 = b$.
- (2) Pour a_n et b_n tel que $f(a_n)f(b_n) < 0$, calculer

$$x_{n+1} = \frac{a_n f(b_n) - b_n f(a_n)}{f(b_n) - f(a_n)}.$$

- (3) Si $f(x_{n+1}) = 0$, arrêter.
- (4) Sinon $f(x_{n+1})$ et $f(a_n)$ sont de signes opposés; prendre $a_{n+1} = a_n$ et $b_{n+1} = x_{n+1}$;
- (5) Sinon prendre $a_{n+1} = x_{n+1}$ et $b_{n+1} = b_n$.
- (6) Répéter (2)–(5) jusqu'à satisfaction du critère d'arrêt (V. la sous-section 7.4.1).

En générale, cette méthode est plus lente que la méthode de Newton. Cependant elle ne requiert pas de dérivée de $f(x)$ et converge toujours vers la racine contenue dans l'intervalle en question. Si l'approche vers la racine est toujours du même côté, on accélère la convergence en remplaçant $f(x)$ du côté stagnant par $f(x)/2$.

TABLE 7.8. Résultats de l'exemple 7.14.

| n | x_n | a_n | b_n | $ x_{n-1} - x_n $ | $f(x_n)$ | $f(a_n)$ |
|-----|----------|----------|-------|-------------------|----------|----------|
| 0 | | 1 | 2 | | | — |
| 1 | 1.333333 | 1.333333 | 2 | | — | — |
| 2 | 1.400000 | 1.400000 | 2 | 0.066667 | — | — |
| 3 | 1.411765 | 1.411765 | 2 | 0.011765 | — | — |
| 4 | 1.413793 | 1.413793 | 2 | 0.002028 | — | — |
| 5 | 1.414141 | 1.414141 | 2 | 0.000348 | — | — |

EXEMPLE 7.14. Approcher $\sqrt{2}$ par la méthode de la position fausse. Arrêter quand $|x_{n+1} - x_n| < 10^{-3}$.

RÉSOLUTION. Il s'agit de trouver la racine positive de l'équation

$$f(x) = x^2 - 2 = 0.$$

On a

$$x_{n+1} = \frac{a_n(b_n^2 - 2) - b_n(a_n^2 - 2)}{(b_n^2 - 2) - (a_n^2 - 2)} = \frac{a_n b_n + 2}{a_n + b_n}.$$

Prendre $a_0 = 1$ et $b_0 = 2$. On voit que $f(1) < 0$ et $f(2) > 0$. Les résultats se trouvent dans le tableau 7.8. Donc $\sqrt{2} \approx 1.414141$. \square

7.5.4. Une méthode globale Newton-bisection. On peut éviter les nombreuses difficultés de la méthode de Newton en la combinant avec la méthode de la bi-section. On profite alors des avantages des 2 méthodes. Au départ, on suppose qu'on a trouvé un intervalle $[a, b]$ qui contient un zéro de $f(x)$, c'est-à-dire $f(a)f(b) < 0$, et que la valeur de départ x_c coïncide avec a ou b . Si

$$x_+ = x_c - \frac{f(x_c)}{f'(x_c)} \in [a, b],$$

on continue avec l'intervalle $[a, x_+]$ ou $[x_+, b]$ qui contient le zéro. Alors on prend $x_c = x_+$. Si le pas de Newton sort de l'intervalle $[a, b]$, on fait un pas avec la bi-section et $x_c = (a + b)/2$. Ordinairement on fait quelques pas avec la bi-section avant de passer à Newton. Cette méthode de Newton globale se trouve dans la fonction M de Matlab qu'on peut télécharger à partir du lien <ftp://ftp.cs.cornell.edu/pub/cv>.

```
function [x,fx,nEvals,aF,bF] = ...
    GlobalNewton(fName,fpName,a,b,tolx,tolf,nEvalsMax)

% Pre:
% fName      string that names a function f(x).
% fpName     string that names the derivative function f'(x).
% a,b       A root of f(x) is sought in the interval [a,b]
%           et f(a)*f(b)<=0.
% tolx,tolf Nonnegative termination criteria.
% nEvalsMax Maximum number of derivative evaluations.
%
% Post:
% x         An approximate zero of f.
```

```

% fx          The value of f at x.
% nEvals     The number of derivative evaluations required.
% aF,bF     The final bracketing interval is [aF,bF].
%
% Comments:
% Iteration terminates as soon as x is within tolX of a true zero
% or if |f(x)| <= tolf or after nEvalMax f-evaluations

fa = feval(fName,a);
fb = feval(fName,b);
if fa*fb>0
    disp('Initial interval not bracketing.')
    return
end
x = a;
fx = feval(fName,x);
fpx = feval(fpName,x);
disp(sprintf('%20.15f %20.15f %20.15f',a,x,b))

nEvals = 1;
while (abs(a-b) > tolX) & (abs(fx) > tolf) &
    ((nEvals<nEvalsMax) | (nEvals==1))
    %[a,b] brackets a root et x = a or x = b.
    if StepIsIn(x,fx,fpx,a,b)
        %Take Newton Step
        disp('Newton')
        x = x-fx/fpx;
    else
        %Take a Bisection Step:
        disp('Bisection')
        x = (a+b)/2;
    end
    fx = feval(fName,x);
    fpx = feval(fpName,x);
    nEvals = nEvals+1;
    if fa*fx<=0
        % There is a root in [a,x]. Bring in right endpoint.
        b = x;
        fb = fx;
    else
        % There is a root in [x,b]. Bring in left endpoint.
        a = x;
        fa = fx;
    end
    disp(sprintf('%20.15f %20.15f %20.15f',a,x,b))
end
aF = a;
bF = b;

```

7.5.5. La fonction `fzero` de Matlab. La fonction `fzero` de MATLAB est une méthode générale pour trouver un zéro d'une fonction sans calcul de dérivées. On appelle la fonction accompagnée d'une valeur de départ x_0 :

`un_zero = fzero('nom_de_la_fonction', x0)`

Le programme produit une valeur près d'un point où la fonction change de signe ou NaN si la recherche flanche. La commande `help fzero` décrit d'autres options.

7.6. Accélération de la convergence d'Aitken-Stefferensen

On peut accélérer la convergence linéaire de la récurrence $x_{n+1} = g(x_n)$ par le processus d'Aitken. Soit une suite $\{x_n\}$ qui converge d'ordre 1 vers le point fixe p . Alors les rapports suivants sont approximativement égaux :

$$\frac{x_{n+1} - p}{x_n - p} \approx \frac{x_{n+2} - p}{x_{n+1} - p}.$$

On en fait une égalité en substituant a_n pour p ,

$$\frac{x_{n+1} - a_n}{x_n - a_n} = \frac{x_{n+2} - a_n}{x_{n+1} - a_n}.$$

On résout pour a_n :

$$a_n = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}.$$

Ceci est le processus d'Aitken qui accélère la convergence dans le sens suivant :

$$\lim_{n \rightarrow \infty} \frac{a_n - p}{x_n - p} = 0.$$

Si l'on introduit la 1ère et la 2ème différences progrades :

$$\Delta x_n = x_{n+1} - x_n,$$

$$\Delta^2 x_n = \Delta(\Delta x_n) = \Delta(x_{n+1} - x_n) = x_{n+2} - 2x_{n+1} + x_n,$$

alors le **processus d'Aitken** devient

$$a_n = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n}. \quad (7.8)$$

Le processus de Stefferensen suppose que $s_1 = a_0$ est une meilleure valeur que x_2 . On obtient donc s_1 par le processus d'Aitken au moyen de $s_0 = x_0$, $z_1 = g(s_0)$ et $z_2 = g(z_1)$. Ensuite, on obtient s_2 par le processus d'Aitken au moyen de s_1 , $z_1 = g(s_1)$ et $z_2 = g(z_2)$. Ainsi de suite. On a l'algorithme suivant :

ALGORITHME 7.4 (Algorithme de Stefferensen). Poser

$$s_0 = x_0,$$

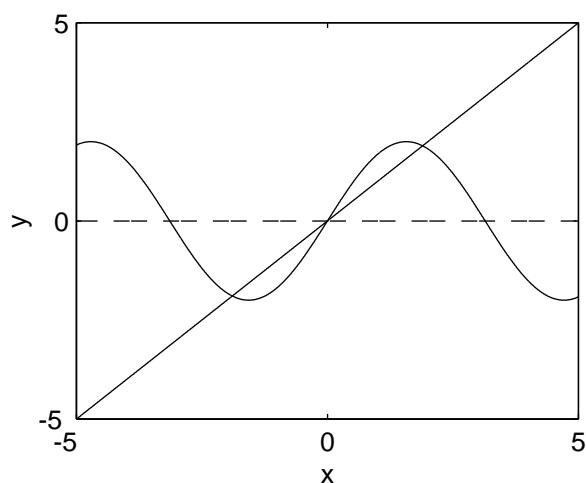
et, pour $n = 0, 1, 2, \dots$,

$$z_1 = g(s_n),$$

$$z_2 = g(z_1),$$

$$s_{n+1} = s_n - \frac{(z_1 - s_n)^2}{z_2 - 2z_1 + s_n}.$$

Le processus de Stefferensen transforme une récurrence d'ordre 1 en une récurrence d'ordre 2.

FIGURE 7.6. Les 3 racines réelles de $x = 2 \sin x$ de l'exemple 7.15.

EXEMPLE 7.15. Soit la récurrence $x_{n+1} = g(x_n)$:

$$x_{n+1} = 2 \sin x_n, \quad x_0 = 1.$$

Itérer 7 fois et appliquer les accélérations d'Aitken et de Steffensen.

RÉSOLUTION. On voit les 3 racines réelles de $x = 2 \sin x$ à la figure 7.6. La fonction `fzero` de Matlab trouve le point fixe près de $x = 1$:

```
p = fzero('x-2*sin(x)',1.)
p = 1.89549426703398
```

La convergence vers p est linéaire puisque

$$g'(p) = -0.63804504828524 \neq 0.$$

La fonction M en Matlab et le script suivants produisent les résultats du tableau 7.9. Les colonnes 2, 3 et 4 contiennent les itérées x_n , et les itérées accélérées a_n et s_n d'Aitken et de Steffensen. La 5ème colonne montre que le rapport

$$\frac{\epsilon_{n+1}}{\epsilon_n^2} = \frac{s_{n+2} - s_{n+1}}{(s_{n+1} - s_n)^2}$$

tend vers une constante. Donc la suite de Steffensen s_n converge d'ordre 2.

La fonction M :

```
function f = twosine(x);
f = 2*sin(x);
```

Le script M :

```
n = 7;
x = ones(1,n+1);
x(1) = 1.0;
for k = 1:n
x(k+1)=twosine(x(k)); % l'iteration x(k+1) = 2*sin(x(k))
end
```

TABLE 7.9. Résultats de l'exemple 7.15.

| n | x_n | a_n | s_n | $\epsilon_{n+1}/\epsilon_n^2$ |
|-----|-------------------|------------------|-------------------|-------------------------------|
| 0 | 1.000000000000000 | 2.23242945471637 | 1.000000000000000 | -0.2620 |
| 1 | 1.68294196961579 | 1.88318435428750 | 2.23242945471637 | 0.3770 |
| 2 | 1.98743653027215 | 1.89201364327283 | 1.83453173271065 | 0.3560 |
| 3 | 1.82890755262358 | 1.89399129067379 | 1.89422502453561 | 0.3689 |
| 4 | 1.93374764234016 | 1.89492839486397 | 1.89549367325365 | 0.3691 |
| 5 | 1.86970615363078 | 1.89525656226218 | 1.89549426703385 | |
| 6 | 1.91131617912526 | | 1.89549426703398 | |
| 7 | 1.88516234821223 | | NaN | |

```

a = ones(1,n-1);
for k = 1:n-1
a(k) = x(k) - (x(k+1)-x(k))^2/(x(k+2)-2*x(k+1)+x(k)); % Aitken
end

s = ones(1,n+1);
s(1) = 1.0;
for k = 1:n
z1=twosine(s(k));
z2=twosine(z1);
s(k+1) = s(k) - (z1-s(k))^2/(z2-2*z1+s(k)); % Steffensen
end

d = ones(1,n-2);
for k = 1:n-2
d(k) = (s(k+2)-s(k+1))/(s(k+1)-s(k))^2; % convergence d'ordre 2
end

```

On voit que le programme de Matlab produit NaN (not a number) pour s_7 à cause de la division par zéro. \square

7.7. La méthode d'Horner et la division synthétique

7.7.1. La méthode de Horner. On minimise le nombre de multiplications dans l'évaluation d'un polynôme en l'exprimant sous forme imbriquée. Par exemple,

$$\begin{aligned}
 p(x) &= a_3x^3 + a_2x^2 + a_1x + a_0 \\
 &= ((a_3x + a_2)x + a_1)x + a_0.
 \end{aligned}$$

Dans ce cas simple, la réduction est de 8 à 3 produits. La commande `horner` de Matlab transforme un polynôme symbolique en sa représentation imbriquée dite de Horner.

```

syms x
p = x^3-6*x^2+11*x-6
p = x^3-6*x^2+11*x-6

```

hp = horner(p)
 hp = -6+(11+(-6+x)*x)*x

La méthode de Horner implémente cette imbrication.

THÉORÈME 7.8 (Méthode de Horner). *Soit*

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Si $b_n = a_n$ et

$$b_k = a_k + b_{k+1} x_0, \quad \text{pour } k = n-1, n-2, \dots, 1, 0,$$

alors

$$b_0 = p(x_0).$$

De plus, si

$$q(x) = b_n x^{n-1} + b_{n-1} x^{n-2} + \cdots + b_2 x + b_1,$$

alors

$$p(x) = (x - x_0)q(x) + b_0.$$

DÉMONSTRATION. Par la définition de $q(x)$,

$$\begin{aligned} (x - x_0)q(x) + b_0 &= (x - x_0)(b_n x^{n-1} + b_{n-1} x^{n-2} + \cdots + b_2 x + b_1) + b_0 \\ &= (b_n x^n + b_{n-1} x^{n-1} + \cdots + b_2 x^2 + b_1 x) \\ &\quad - (b_n x_0 x^{n-1} + b_{n-1} x_0 x^{n-2} + \cdots + b_2 x_0 x + b_1 x_0) + b_0 \\ &= b_n x^n + (b_{n-1} - b_n x_0) x^{n-1} + \cdots + (b_1 - b_2 x_0) x + (b_0 - b_1 x_0) \\ &= a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \\ &= p(x) \end{aligned}$$

et

$$b_0 = p(x_0). \quad \square$$

7.7.2. La division synthétique. L'évaluation d'un polynôme en $x = x_0$ par la méthode de Horner revient à en faire la division synthétique comme le montre l'exemple 7.16.

EXEMPLE 7.16. Évaluer le polynôme

$$p(x) = 2x^4 - 3x^2 + 3x - 4$$

en $x_0 = -2$ par la méthode de Horner.

RÉSOLUTION. En multipliant successivement les éléments de la 3ème ligne du tableau suivant par $x_0 = -2$ et en additionnant les résultats à la 1ère ligne on obtient $p(-2)$.

| | | | | |
|-----------|------------|------------|------------|------------|
| $a_4 = 2$ | $a_3 = 0$ | $a_2 = -3$ | $a_1 = 3$ | $a_0 = -4$ |
| | -4 | 8 | -10 | 14 |
| $b_4 = 2$ | $b_3 = -4$ | $b_2 = 5$ | $b_1 = -7$ | $b_0 = 10$ |

Alors

$$p(x) = (x + 2)(2x^3 - 4x^2 + 5x - 7) + 10$$

et

$$p(-2) = 10. \quad \square$$

La méthode de Horner est très efficace pour trouver les zéros d'un polynôme $p(x)$ par la méthode de Newton. Soit

$$p(x) = (x - x_0)q(x) + b_0$$

et sa dérivée

$$p'(x) = (x - x_0)q'(x) + q(x).$$

Alors

$$p'(x_0) = q(x_0).$$

En substituant ces résultats dans la méthode de Newton, on a

$$\begin{aligned} x_n &= x_{n-1} - \frac{p(x_{n-1})}{p'(x_{n-1})} \\ &= x_{n-1} - \frac{p(x_{n-1})}{q(x_{n-1})}. \end{aligned}$$

On applique cette procédure à l'exemple 7.17.

EXEMPLE 7.17. Calculer la valeur du polynôme

$$p(x) = 2x^4 - 3x^2 + 3x - 4$$

et de sa dérivée $p'(x)$ en $x_0 = -2$ par la méthode de Horner et employer ces résultats pour trouver la 1ère itérée x_1 de la méthode de Newton.

RÉSOLUTION. En multipliant successivement les éléments de la 3ème ligne du tableau suivant par $x_0 = -2$ et en additionnant les résultats à la 1ère ligne on obtient $p(-2)$. Ensuite, en multipliant successivement les éléments de la 5ème ligne du tableau suivant par $x_0 = -2$ et en additionnant les résultats à la 3ème ligne on obtient $p'(-2)$.

| | | | | | |
|---|----|----|----------------|--------------|--|
| 2 | 0 | -3 | 3 | -4 | |
| | -4 | 8 | -10 | 14 | |
| | ↗ | ↗ | ↗ | ↗ | |
| 2 | -4 | 5 | -7 | 10 = $p(-2)$ | |
| | -4 | 16 | -42 | | |
| | ↗ | ↗ | ↗ | | |
| 2 | -8 | 21 | -49 = $p'(-2)$ | | |

Alors

$$p(-2) = 10, \quad p'(-2) = -49,$$

et

$$x_1 = -2 - \frac{10}{-49} \approx -1.7959. \quad \square$$

7.8. La méthode de Müller

La méthode de Müller, ou méthode de la parabole, trouve les zéros réels ou complexes d'une équation

$$f(x) = 0.$$

Cette méthode emploie 3 approximations initiales, x_0 , x_1 et x_2 , pour construire le parabole

$$p(x) = a(x - x_2)^2 + b(x - x_2) + c,$$

qui passe par les 3 points $(x_0, f(x_0))$, $(x_1, f(x_1))$ et $(x_2, f(x_2))$ de la courbe $f(x)$. La prochaine approximation x_3 est le point d'intersection de la parabole avec l'axe des réels le plus près de x_2 .

On obtient les coefficients a , b et c de la parabole en résolvant le système linéaire

$$\begin{aligned} f(x_0) &= a(x_0 - x_2)^2 + b(x_0 - x_2) + c, \\ f(x_1) &= a(x_1 - x_2)^2 + b(x_1 - x_2) + c, \\ f(x_2) &= c. \end{aligned}$$

On a immédiatement

$$c = f(x_2)$$

et l'on obtient a et b à partir du système linéaire

$$\begin{bmatrix} (x_0 - x_2)^2 & (x_0 - x_2) \\ (x_1 - x_2)^2 & (x_1 - x_2) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} f(x_0) - f(x_2) \\ f(x_1) - f(x_2) \end{bmatrix}.$$

Alors, on pose

$$p(x_3) = a(x_3 - x_2)^2 + b(x_3 - x_2) + c = 0$$

qu'on résout pour $x_3 - x_2$:

$$\begin{aligned} x_3 - x_2 &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \times \frac{-b \mp \sqrt{b^2 - 4ac}}{-b \mp \sqrt{b^2 - 4ac}} \\ &= \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}. \end{aligned}$$

Pour trouver x_3 le plus près de x_2 , on maximise le dénominateur:

$$x_3 = x_2 - \frac{2c}{b + \text{sign}(b)\sqrt{b^2 - 4ac}}.$$

L'ordre de convergence de la méthode de Müller vers une racine simple ou double est approximativement 1.839. La convergence vers une racine triple n'est pas assurée.

EXEMPLE 7.18. Trouver par la méthode de Müller les 4 zéros du polynôme

$$16x^4 - 40x^3 + 5x^2 + 20x + 6$$

tracé dans la Fig. 7.7.

RÉSOLUTION. Les commandes de Matlab suivantes font une passe de la méthode de Müller sur le polynôme donné écrit sous forme imbriquée.

```
syms x
pp = 16*x^4-40*x^3+5*x^2+20*x+6
pp = 16*x^4-40*x^3+5*x^2+20*x+6
pp = horner(pp)
pp = 6+(20+(5+(-40+16*x)*x)*x)*x
```

Le polyôme est évalué par la fonction M de Matlab:

```
function pp = mullerpol(x);
pp = 6+(20+(5+(-40+16*x)*x)*x)*x;
```

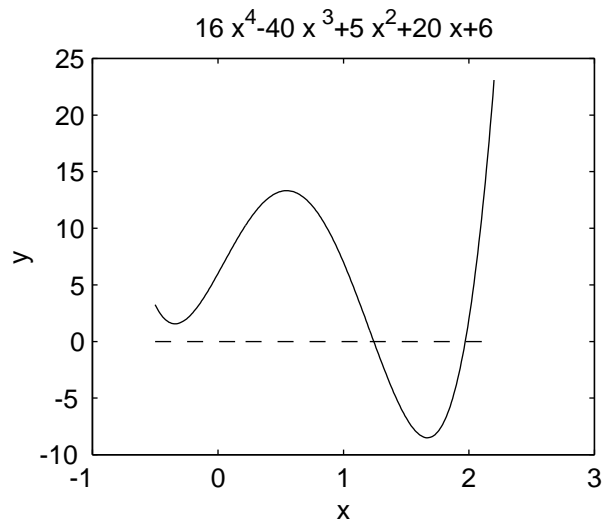


FIGURE 7.7. Le tracé du polynôme $16x^4 - 40x^3 + 5x^2 + 20x + 6$ de l'exemple 7.18.

La méthode de Müller calcule x_3 au moyen des 3 valeurs initiales données:

```
x0 = 0.5; x1 = -0.5; x2 = 0; % starting values
m = [(x0-x2)^2 x0-x2; (x1-x2)^2 x1-x2];
rhs = [mullerpol(x0)-mullerpol(x2); mullerpol(x1)-mullerpol(x2)];
ab = m\rhs; a = ab(1); b = ab(2); % coefficients a et b
c = mullerpol(x2); % coefficient c
x3 = x2 -(2*c)/(b+sign(b)*sqrt(b^2-4*a*c))
x3 = -0.5556 + 0.5984i
```

On itère la méthode jusqu'à la convergence. Les 4 zéros sont :

```
rr = roots([16 -40 5 20 6])'
rr = 1.9704 1.2417 -0.3561 - 0.1628i -0.3561 + 0.1628i
```

On peut obtenir les zéros réels par la méthode de Müller avec les valeurs de départ: $[0.5, 1.0, 1.5]$ et $[2.5, 2.0, 2.25]$. \square

Interpolation et extrapolation

Dans les applications on ne connaît souvent que quelques valeurs discrètes d'une fonction inconnue $f(x)$ fournies par les résultats expérimentaux :

$$(x_0, f_0), \quad (x_1, f_1), \quad (x_2, f_2), \quad \dots, \quad (x_n, f_n), \quad (8.1)$$

où les f_i dénotent les valeurs observées $f(x_i)$. On veut employer ces données pour approcher $f(x)$ en un point $x \neq x_i$ quelconque.

En terme technique on parle d'**interpolation** de $f(x)$ en x si le point x est entre deux valeurs des x_i , sinon on parle d'**extrapolation** de $f(x)$ en x . Dans les 2 cas, on construit un polynôme d'interpolation $p_n(x)$ de degré n qui passe par les $n + 1$ points énumérés en (8.1). Ce polynôme servira à estimer $f(x)$.

8.1. Polynômes d'interpolation de Lagrange

On exprime le polynôme d'interpolation de Lagrange $p_n(x)$ de degré n qui passe par les $n + 1$ points $(x_k, f(x_k))$, $k = 0, 1, \dots, n$, au moyen de la base de Lagrange:

$$L_k(x) = \frac{(x - x_0)(x - x_1) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_n)}{(x_k - x_0)(x_k - x_1) \cdots (x_k - x_{k-1})(x_k - x_{k+1}) \cdots (x_k - x_n)}.$$

On voit que $L_k(x)$ est un polynôme de degré n et que

$$L_k(x) = \begin{cases} 1, & x = x_k, \\ 0, & x = x_j, \quad j \neq k. \end{cases}$$

Alors le **polynôme d'interpolation de Lagrange** de $f(x)$ est

$$p_n(x) = f(x_0)L_0(x) + f(x_1)L_1(x) + \cdots + f(x_n)L_n(x). \quad (8.2)$$

Il est de degré n et interpole $f(x)$ en les points énumérés en (8.1).

EXEMPLE 8.1. Interpoler $f(x) = 1/x$ aux nœuds $x_0 = 2$, $x_1 = 2.5$ et $x_2 = 4$ par un polynôme d'interpolation de Lagrange de degré 2.

RÉSOLUTION. La base de Lagrange sous forme imbriquée est

$$\begin{aligned} L_0(x) &= \frac{(x - 2.5)(x - 4)}{(2 - 2.5)(2 - 4)} = (x - 6.5)x + 10, \\ L_1(x) &= \frac{(x - 2)(x - 4)}{(2.5 - 2)(2.5 - 4)} = \frac{(-4x + 24)x - 32}{3}, \\ L_2(x) &= \frac{(x - 2)(x - 2.5)}{(4 - 2)(4 - 2.5)} = \frac{(x - 4.5)x + 5}{3}. \end{aligned}$$

Donc

$$\begin{aligned} p(x) &= \frac{1}{2} [(x - 6.5)x + 10] + \frac{1}{2.5} \frac{(-4x + 24)x - 32}{3} + \frac{1}{4} \frac{(x - 4.5)x + 5}{3} \\ &= (0.05x - 0.425)x + 1.15. \quad \square \end{aligned}$$

THÉORÈME 8.1. Soit x_0, x_1, \dots, x_n , $n + 1$ points distincts dans l'intervalle $[a, b]$ et $f \in C^{n+1}[a, b]$. Alors il existe un nombre $\xi(x) \in [a, b]$ tel que

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} (x - x_0)(x - x_1) \cdots (x - x_n), \quad (8.3)$$

où $p_n(x)$ est le polynôme d'interpolation de Lagrange. En particulier, si

$$m_{n+1} = \min_{a \leq x \leq b} |f^{(n+1)}(x)| \quad \text{et} \quad M_{n+1} = \max_{a \leq x \leq b} |f^{(n+1)}(x)|,$$

alors l'erreur absolue en $p_n(x)$ est bornée par les inégalités:

$$\begin{aligned} \frac{m_{n+1}}{(n+1)!} |(x - x_0)(x - x_1) \cdots (x - x_n)| &\leq |f(x) - p_n(x)| \\ &\leq \frac{M_{n+1}}{(n+1)!} |(x - x_0)(x - x_1) \cdots (x - x_n)| \end{aligned}$$

pour tout $a \leq x \leq b$.

PROOF. En premier, on voit que l'erreur est 0 en $x = x_0, x_1, \dots, x_n$ puisque

$$p_n(x_k) = f(x_k), \quad k = 0, 1, \dots, n,$$

d'après la propriété d'interpolation de $p_n(x)$. Pour $x \neq x_k$ quelconque, on définit la fonction auxiliaire

$$\begin{aligned} g(t) &= f(t) - p_n(t) - [f(x) - p_n(x)] \frac{(t - x_0)(t - x_1) \cdots (t - x_n)}{(x - x_0)(x - x_1) \cdots (x - x_n)} \\ &= f(t) - p_n(t) - [f(x) - p_n(x)] \prod_{i=0}^n \frac{t - x_i}{x - x_i}. \end{aligned}$$

En $t = x_k$,

$$g(x_k) = f(x_k) - p_n(x_k) - [f(x) - p_n(x)] \times 0 = 0$$

et en $t = x$,

$$g(x) = f(x) - p_n(x) - [f(x) - p_n(x)] \times 1 = 0.$$

Alors $g \in C^{n+1}[a, b]$ admet $n + 2$ zéros sur $[a, b]$. Par le théorème de Rolle généralisé, $g'(t)$ admet $n + 1$ zéros sur $[a, b]$, $g''(t)$ admet n zéros sur $[a, b]$, \dots , $g^{(n+1)}(t)$ admet 1 zéro, $\xi \in [a, b]$:

$$\begin{aligned} g^{(n+1)}(\xi) &= f^{(n+1)}(\xi) - p_n^{(n+1)}(\xi) - [f(x) - p_n(x)] \frac{d^{n+1}}{dt^{n+1}} \left[\prod_{i=0}^n \frac{t - x_i}{x - x_i} \right]_{t=\xi} \\ &= f^{(n+1)}(\xi) - 0 - [f(x) - p_n(x)] \frac{(n+1)!}{\prod_{i=0}^n (x - x_i)} \\ &= 0 \end{aligned}$$

puisque $p_n(x)$ est un polynôme de degré n de sorte que sa $(n+1)$ ème dérivée est nulle et seulement le 1er terme, t^{n+1} , du produit $\prod_{i=0}^n (t-x_i)$ contribue $(n+1)!$ dans la $(n+1)$ ème dérivée. Donc

$$f(x) = p_n(x) + \frac{f^{(n+1)}(\xi(x))}{(n+1)!} (x-x_0)(x-x_1)\cdots(x-x_n). \quad \square$$

La représentation (8.2) de $p_n(x)$ est coûteuse en calculs et doit être refaite si l'on veut augmenter le degré de $p_n(x)$ pour une interpolation plus fine.

Si les points x_i sont distincts, le polynôme est unique. En effet, si $p_n(x)$ et $q_n(x)$ de degré n interpolent $f(x)$ en $n+1$ points distincts, alors

$$p_n(x) - q_n(x)$$

est un polynôme de degré n qui admet $n+1$ zéros distincts. Il est donc identiquement nul.

8.2. Interpolation newtonnienne aux différences divisées

On représente les polynômes d'interpolation newtonniens aux différences divisées, $p_n(x)$, de degré n au moyen d'une base factorielle :

$$p_n(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \cdots + a_n(x-x_0)(x-x_1)\cdots(x-x_{n-1}).$$

On détermine les coefficients a_k par récurrence. On note

$$f_k = f(x_k).$$

Soit $x_0 \neq x_1$. On considère les 2 points: (x_0, f_0) et (x_1, f_1) . La propriété d'interpolation du polynôme

$$p_1(x) = a_0 + a_1(x-x_0)$$

implique que

$$p_1(x_0) = a_0 = f_0, \quad p_1(x_1) = f_0 + a_1(x_1-x_0) = f_1.$$

On résout pour a_1 :

$$a_1 = \frac{f_1 - f_0}{x_1 - x_0}.$$

On note

$$f[x_0, x_1] = \frac{f_1 - f_0}{x_1 - x_0}$$

la **1ère différence divisée**. Alors, le polynôme d'interpolation aux différences divisées du 1er degré est

$$p_1(x) = f_0 + (x-x_0)f[x_0, x_1].$$

EXEMPLE 8.2. Soit une fonction $f(x)$ passant par les 2 points (2.2, 6.2) et (2.5, 6.7). Trouver le polynôme d'interpolation aux différences divisées du 1er degré $p_1(x)$ de $f(x)$ et interpoler f en $x = 2.35$ au moyen de $p_1(x)$.

RÉSOLUTION. Puisque

$$f[2.2, 2.5] = \frac{6.7 - 6.2}{2.5 - 2.2} = 1.6667,$$

alors

$$p_1(x) = 6.2 + (x-2.2) \times 1.6667 = 2.5333 + 1.6667x.$$

En particulier, $p_1(2.35) = 6.45$. □

EXEMPLE 8.3. Interpoler $\cos 0.2$ par un polynôme du 1er degré au moyen des valeurs $\cos 0$ et $\cos \pi/8$.

RÉSOLUTION. On a les 2 points :

$$(0, \cos 0) = (0, 1) \quad \text{and} \quad \left(\frac{\pi}{8}, \cos \frac{\pi}{8}\right) = \left(\frac{\pi}{8}, \frac{1}{2}\sqrt{\sqrt{2}+2}\right)$$

(Si l'on substitue $\theta = \pi/8$ dans la formule

$$\cos^2 \theta = \frac{1 + \cos(2\theta)}{2}$$

on obtient

$$\cos \frac{\pi}{8} = \frac{1}{2}\sqrt{\sqrt{2}+2}$$

puisque $\cos(\pi/4) = \sqrt{2}/2$.) Alors,

$$f[0, \pi/8] = \frac{\left(\frac{\sqrt{\sqrt{2}+2}}{2} - 1\right)}{\pi/8 - 0} = \frac{4}{\pi} \left(\sqrt{\sqrt{2}+2} - 2\right)$$

et l'on obtient

$$p_1(x) = 1 + \frac{4}{\pi} \left(\sqrt{\sqrt{2}+2} - 2\right) x.$$

En particulier,

$$p_1(0.2) = 0.96125.$$

On remarque que $\cos 0.2 = 0.98007$ (arrondi à 5 décimales). L'erreur absolue est 0.01882. \square

Soit les 3 points

$$(x_0, f_0), \quad (x_1, f_1), \quad (x_2, f_2), \quad \text{où } x_i \neq x_j \text{ pour } i \neq j.$$

Le polynôme d'interpolation aux différences divisées de degré 2 qui passe par ces points est

$$p_2(x) = f_0 + (x - x_0) f[x_0, x_1] + (x - x_0)(x - x_1) f[x_0, x_1, x_2]$$

où

$$f[x_0, x_1] := \frac{f_1 - f_0}{x_1 - x_0} \quad \text{et} \quad f[x_0, x_1, x_2] := \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

sont les différences divisées du **premier** et du **second** degré.

EXEMPLE 8.4. Interpoler une fonction $f(x)$ en les 3 points

$$(2.2, 6.2), \quad (2.5, 6.7), \quad (2.7, 6.5),$$

au moyen d'un polynôme d'interpolation aux différences divisées $p_2(x)$ de degré 2 et interpoler $f(x)$ en $x = 2.35$ par la valeur $p_2(2.35)$.

RÉSOLUTION. On a

$$f[2.2, 2.5] = 1.6667, \quad f[2.5, 2.7] = -1$$

et

$$f[2.2, 2.5, 2.7] = \frac{f[2.5, 2.7] - f[2.2, 2.5]}{2.7 - 2.2} = \frac{-1 - 1.6667}{2.7 - 2.2} = -5.3334.$$

Alors,

$$p_2(x) = 6.2 + (x - 2.2) \times 1.6667 + (x - 2.2)(x - 2.5) \times (-5.3334).$$

En particulier, $p_2(2.35) = 6.57$. \square

EXEMPLE 8.5. Intepoler $\cos x$ par le polynôme d'interpolation aux différences divisées de degré 2 au moyen des valeurs $\cos 0$, $\cos \pi/8$ et $\cos \pi/4$, et approcher $\cos 0.2$.

RÉSOLUTION. On a vu à l'exemple 8.3 que

$$\cos \frac{\pi}{8} = \frac{1}{2} \sqrt{\sqrt{2} + 2}.$$

Donc, à partir des 3 points

$$(0, 1), \quad (\pi/8, \cos \pi/8), \quad (\pi/4, \sqrt{2}/2),$$

on obtient les différences divisées

$$f[0, \pi/8] = \frac{4}{\pi} \left(\sqrt{\sqrt{2} + 2} - 2 \right), \quad f[\pi/8, \pi/4] = \frac{4}{\pi} \left(\sqrt{2} - \sqrt{\sqrt{2} + 2} \right),$$

et

$$\begin{aligned} f[0, \pi/8, \pi/4] &= \frac{f[\pi/8, \pi/4] - f[0, \pi/8]}{\pi/4 - 0} \\ &= \frac{4}{\pi} \left[\frac{\sqrt{2}/2 - (\sqrt{\sqrt{2} + 2})/2}{\pi/4 - \pi/8} - \frac{4\sqrt{\sqrt{2} + 2} - 8}{\pi} \right] \\ &= \frac{16}{\pi^2} \left(\sqrt{2} - 2\sqrt{\sqrt{2} + 2} \right). \end{aligned}$$

Donc,

$$p_2(x) = 1 + x \frac{4}{\pi} \left(\sqrt{\sqrt{2} + 2} - 2 \right) + x \left(x - \frac{\pi}{8} \right) \frac{16}{\pi^2} \left(\sqrt{2} - 2\sqrt{\sqrt{2} + 2} \right).$$

Enfin,

$$p_2(0.2) = 0.97881.$$

L'erreur absolue est 0.00189. □

Soit $n + 1$ donnés quelconques

$$(x_0, f_0), \quad (x_1, f_1), \quad \dots, \quad (x_n, f_n),$$

où $x_i \neq x_j$ for $i \neq j$. Le polynôme d'interpolation newtonien aux différences divisées de degré n est

$$\begin{aligned} p_n(x) &= f_0 + (x - x_0) f[x_0, x_1] + (x - x_0)(x - x_1) f[x_0, x_1, x_2] + \dots \\ &\quad + (x - x_0)(x - x_1) \dots (x - x_{n-1}) f[x_0, x_1, \dots, x_n], \end{aligned} \quad (8.4)$$

où, par définition,

$$f[x_j, x_{j+1}, \dots, x_k] = \frac{f[x_{j+1}, \dots, x_k] - f[x_j, x_{j+1}, \dots, x_{k-1}]}{x_k - x_j}$$

est la $(k - j)$ ème différence divisée. On obtient cette formule par récurrence.

Le tableau 8.1 contient une table de différences divisées.

EXEMPLE 8.6. Construire le polynôme d'interpolation passant par les 4 points inégalement espacés :

$$(1.0, 2.4), \quad (1.3, 2.2), \quad (1.5, 2.3), \quad (1.7, 2.4),$$

sur le tracé d'une certaine fonction $f(x)$. Approcher $f(1.4)$.

TABLE 8.1. Table de différences divisées

| x | $f(x)$ | 1ères différences divisées | 2èmes différences divisées | 3èmes différences divisées |
|-------|----------|-------------------------------|-------------------------------|-------------------------------|
| x_0 | $f[x_0]$ | | | |
| x_1 | $f[x_1]$ | $f[x_0, x_1]$ | | |
| x_2 | $f[x_2]$ | $f[x_1, x_2]$ | $f[x_0, x_1, x_2]$ | |
| x_3 | $f[x_3]$ | $f[x_2, x_3]$ | $f[x_1, x_2, x_3]$ | $f[x_0, x_1, x_2, x_3]$ |
| x_4 | $f[x_4]$ | $f[x_3, x_4]$ | $f[x_2, x_3, x_4]$ | $f[x_1, x_2, x_3, x_4]$ |
| x_5 | $f[x_5]$ | $f[x_4, x_5]$ | $f[x_3, x_4, x_5]$ | $f[x_2, x_3, x_4, x_5]$ |

RÉSOLUTION. On construit d'abord la table des différences divisées de Newton :

| | | | | |
|-------|----------|-------------------|----------------------------|-------------------------------------|
| x_i | $f(x_i)$ | $f[x_i, x_{i+1}]$ | $f[x_i, x_{i+1}, x_{i+2}]$ | $f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$ |
| 1.0 | 2.4 | | | |
| | | -0.66667 | | |
| 1.3 | 2.2 | | 2.33333 | |
| | | 0.500000 | | -3.33333 |
| 1.5 | 2.3 | | 0.00000 | |
| | | 0.500000 | | |
| 1.7 | 2.4 | | | |

On obtient alors le polynôme :

$$p_3(x) = 2.4 + (x - 1.0)(-0.66667) + (x - 1.0)(x - 1.3) \times 2.33333 \\ + (x - 1.0)(x - 1.3)(x - 1.5)(-3.33333).$$

L'approximation de $f(1.4)$ est

$$p_3(1.4) = 2.2400. \quad \square$$

8.3. Interpolation prograde de Gregory–Newton

On récrit (8.4) dans le cas spécial de nœuds équidistants,

$$x_i = x_0 + i h.$$

La **1ère** et la **2ème différences progrades** de $f(x)$ en x_j sont

$$\Delta f_j := f_{j+1} - f_j, \quad \Delta^2 f_j := \Delta f_{j+1} - \Delta f_j.$$

En général la **k ème différence prograde** de $f(x)$ en x_j est

$$\Delta^k f_j := \Delta^{k-1} f_{j+1} - \Delta^{k-1} f_j.$$

Par induction mathématique on obtient

$$f[x_0, \dots, x_k] := \frac{1}{k! h^k} \Delta^k f_0.$$

Si l'on pose

$$r = \frac{x - x_0}{h},$$

alors, on obtient, pour des nœuds équidistants,

$$x - x_k = x - x_0 - (x_k - x_0) = hr - hk = h(r - k)$$

et

$$(x - x_0)(x - x_1) \cdots (x - x_{k-1}) = h^k r(r - 1)(r - 2) \cdots (r - k + 1).$$

Donc la formule (8.4) devient

$$\begin{aligned} p_n(r) &= f_0 + \sum_{k=1}^n \frac{r(r-1) \cdots (r-k+1)}{k!} \Delta^k f_0 \\ &= \sum_{k=0}^n \binom{r}{k} \Delta^k f_0, \end{aligned} \tag{8.5}$$

où

$$\binom{r}{k} = \begin{cases} \frac{r(r-1) \cdots (r-k+1)}{k!}, & \text{if } k > 0, \\ 1, & \text{if } k = 0. \end{cases}$$

Le polynôme (8.5) est le **polynôme d'interpolation prograde de Gregory-Newton**.

EXEMPLE 8.7. Soit les données équidistantes :

| | | | | | | |
|-----|-------|-------|-------|-------|-------|-------|
| x | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 |
| y | 35000 | 36000 | 36500 | 37000 | 37800 | 39000 |

Extrapoler y en 1994.

RÉSOLUTION. On construit la table de différences progrades :

| i | x_i | y_i | Δy_i | $\Delta^2 y_i$ | $\Delta^3 y_i$ | $\Delta^4 y_i$ | $\Delta^5 y_i$ |
|-----|-------|-------|--------------|----------------|----------------|----------------|----------------|
| 0 | 1988 | 35000 | 1000 | | | | |
| 1 | 1989 | 36000 | 500 | -500 | | | |
| 2 | 1990 | 36500 | 500 | 0 | 500 | -200 | |
| 3 | 1991 | 37000 | 800 | 300 | 100 | -200 | 0 |
| 4 | 1992 | 37800 | 1200 | 400 | | | |
| 5 | 1993 | 39000 | | | | | |

Si l'on pose $r = (x - 1988)/1$, alors :

$$\begin{aligned} p_5(r) &= 35000 + r(1000) + \frac{r(r-1)}{2}(-500) + \frac{r(r-1)(r-2)}{6}(500) \\ &\quad + \frac{r(r-1)(r-2)(r-3)}{24}(-200) + \frac{r(r-1)(r-2)(r-3)(r-4)}{120}(0). \end{aligned}$$

L'extrapolation des données en 1994 avec $r = 6$ donne

$$p_5(6) = 40500.$$

L'itération de la commande `diff(y,n)` de Matlab produit la table de différences suivante:

```

y = [35000 36000 36500 37000 37800 39000]
dy = diff(y);
    dy = 1000          500          500          800          1200
d2y = diff(y,2)
    d2y = -500      0      300      400
d3y = diff(y,3)
    d3y = 500      300      100
d4y = diff(y,4)
    d4y = -200     -200
d5y = diff(y,5)
    d5y = 0

```

□

EXEMPLE 8.8. Approcher $f(1.5)$ au moyen des données équidistantes :

| | | | | | |
|--------|-----------|-----------|-----------|-----------|-----------|
| x | 1.0 | 1.3 | 1.6 | 1.9 | 2.2 |
| $f(x)$ | 0.7651977 | 0.6200860 | 0.4554022 | 0.2818186 | 0.1103623 |

RÉSOLUTION. On construit la table de différences progrades :

| i | x_i | y_i | Δy_i | $\Delta^2 y_i$ | $\Delta^3 y_i$ | $\Delta^4 y_i$ |
|-----|-------|-----------|--------------|----------------|----------------|----------------|
| 0 | 1.0 | 0.7651977 | | | | |
| 1 | 1.3 | 0.6200860 | -0.145112 | | | |
| 2 | 1.6 | 0.4554022 | -0.164684 | -0.0195721 | | |
| 3 | 1.9 | 0.2818186 | -0.173584 | -0.0088998 | 0.0106723 | |
| 4 | 2.2 | 0.1103623 | -0.170856 | 0.0021273 | 0.0110271 | 0.0003548 |

Si l'on pose $r = (x - 1.0)/0.3$, on obtient

$$\begin{aligned}
 p_4(r) = & 0.7651977 + r(-0.145112) + \frac{r(r-1)}{2}(-0.0195721) \\
 & + \frac{r(r-1)(r-2)}{6}(0.0106723) + \frac{r(r-1)(r-2)(r-3)}{24}(0.0003548).
 \end{aligned}$$

Pour interpolier $f(x)$ en $x = 1.5$, on prend $r = 5/3$ et l'on obtient

$$p_4(5/3) = 0.511819. \quad \square$$

8.4. Polynômes d'interpolation rétrograde de Gregory-Newton

On emploie un polynôme d'interpolation aux différences rétrogrades de Gregory-Newton pour interpoler une fonction en une valeur près du bas d'une table de différences contruite sur les données

$$(x_{-n}, f_{-n}), (x_{-n+1}, f_{-n+1}), \dots, (x_0, f_0).$$

On pose

$$r = \frac{x - x_0}{h}.$$

Alors, pour des nœuds équidistants, on a

$$x - x_{-k} = x - x_0 - (x_{-k} - x_0) = hr + hk = h(r + k)$$

et

$$(x - x_0)(x - x_{-1}) \cdots (x - x_{-(k-1)}) = h^k r(r+1)(r+2) \cdots (r+k-1).$$

Donc (8.5) devient

$$\begin{aligned} p_n(r) &= f_0 + \sum_{k=1}^n \frac{r(r+1) \cdots (r+k-1)}{k!} \Delta^k f_{-k} \\ &= \sum_{k=0}^n \binom{r+k-1}{k} \Delta^k f_{-k}, \end{aligned} \quad (8.6)$$

Le polynôme (8.6) est le **polynôme d'interpolation aux différences rétrogrades de Gregory-Newton**.

EXEMPLE 8.9. Interpoler les données équidistantes de l'exemple 8.8 en $x = 2.1$.

RÉSOLUTION. On construit la table de différences :

| i | x_i | y_i | Δy_i | $\Delta^2 y_i$ | $\Delta^3 y_i$ | $\Delta^4 y_i$ |
|-----|-------|-----------|--------------|----------------|----------------|----------------|
| -4 | 1.0 | 0.7651977 | | | | |
| | | | -0.145112 | | | |
| -3 | 1.3 | 0.6200860 | | -0.0195721 | | |
| | | | -0.164684 | | 0.0106723 | |
| -2 | 1.6 | 0.4554022 | | -0.0088998 | | 0.0003548 |
| | | | -0.173584 | | 0.0110271 | |
| -1 | 1.9 | 0.2818186 | | 0.0021273 | | |
| | | | -0.170856 | | | |
| 0 | 2.2 | 0.1103623 | | | | |

Si l'on pose $r = (x - 2.2)/0.3$, on a

$$\begin{aligned} p_4(r) &= 0.1103623 + r(-0.170856) + \frac{r(r+1)}{2} (0.0021273) \\ &\quad + \frac{r(r+1)(r+2)}{6} (0.0110271) + \frac{r(r+1)(r+2)(r+3)}{24} (0.0003548). \end{aligned}$$

Puisque

$$r = \frac{2.1 - 2.2}{0.3} = -\frac{1}{3},$$

alors

$$p_4(-1/3) = 0.115904. \quad \square$$

8.5. Polynômes d'interpolation d'Hermite

Soit $n + 1$ nœuds distincts x_0, x_1, \dots, x_n et $2n + 2$ valeurs $f_k = f(x_k)$ et $f'_k = f'(x_k)$. Le polynôme d'interpolation d'Hermite $p_{2n+1}(x)$ de degré $2n + 1$,

$$p_{2n+1}(x) = \sum_{m=0}^n h_m(x) f_m + \sum_{m=0}^n \hat{h}_m(x) f'_m,$$

prend les valeurs

$$p_{2n+1}(x_k) = f_k, \quad p'_{2n+1}(x_k) = f'_k, \quad k = 0, 1, \dots, n.$$

On construit des polynômes $h_m(x)$ et $\hat{h}_m(x)$ de degré au plus $2n + 1$ qui satisfont les conditions suivante :

$$\begin{aligned} h_m(x_k) &= h'_m(x_k) = 0, & k \neq m, \\ h_m(x_m) &= 1, \\ h'_m(x_m) &= 0, \end{aligned}$$

et

$$\begin{aligned} \hat{h}_m(x_k) &= \hat{h}'_m(x_k) = 0, & k \neq m, \\ \hat{h}_m(x_m) &= 0, \\ \hat{h}'_m(x_m) &= 1. \end{aligned}$$

Ces conditions sont satisfaites par les polynômes

$$h_m(x) = [1 - 2(x - x_m)L'_m(x_m)]L_m^2(x)$$

et

$$\hat{h}_m(x) = (x - x_m)L_m^2(x),$$

où

$$L_m(x) = \prod_{k=0, k \neq m}^n \frac{x - x_k}{x_m - x_k}$$

sont les éléments de la base de Lagrange de degré n .

Une méthode pratique pour construire les polynômes d'interpolation d'Hermite sur $n + 1$ nœuds distincts x_0, x_1, \dots, x_n consiste à poser

$$z_{2i} = z_{2i+1} = x_i, \quad i = 0, 1, \dots, n,$$

et à prendre

$$f'(x_0) \text{ pour } f[z_0, z_1], \quad f'(x_1) \text{ pour } f[z_2, z_3], \quad \dots, \quad f'(x_j) \text{ pour } f[z_{2n}, z_{2n+1}]$$

dans la table de différences divisées pour un polynôme d'interpolation d'Hermite de degré $2n + 1$. Alors,

$$p_{2n+1}(x) = f[z_0] + \sum_{k=1}^{2n+1} f[z_0, z_1, \dots, z_k](x - z_0)(x - z_1) \cdots (x - z_{k-1}).$$

Voici une table de différences divisées pour un polynôme d'interpolation d'Hermite :

| z | $f(z)$ | 1ères différences divisées | 2èmes différences divisées | 3èmes différences divisées |
|-------------|-------------------|-------------------------------|-------------------------------|-------------------------------|
| $z_0 = x_0$ | $f[z_0] = f(x_0)$ | | | |
| | | $f[z_0, z_1] = f'(x_0)$ | | |
| $z_1 = x_0$ | $f[z_1] = f(x_0)$ | | $f[z_0, z_1, z_2]$ | |
| | | $f[z_1, z_2]$ | | $f[z_0, z_1, z_2, z_3]$ |
| $z_2 = x_1$ | $f[z_2] = f(x_1)$ | | $f[z_1, z_2, z_3]$ | |
| | | $f[z_2, z_3] = f'(x_1)$ | | $f[z_1, z_2, z_3, z_4]$ |
| $z_3 = x_1$ | $f[z_3] = f(x_1)$ | | $f[z_2, z_3, z_4]$ | |
| | | $f[z_3, z_4]$ | | $f[z_2, z_3, z_4, z_5]$ |
| $z_4 = x_2$ | $f[z_4] = f(x_2)$ | | $f[z_3, z_4, z_5]$ | |
| | | $f[z_4, z_5] = f'(x_2)$ | | |
| $z_5 = x_2$ | $f[z_5] = f(x_2)$ | | | |

EXEMPLE 8.10. Interpoler en $x = 1.5$ les données soulignées de la table suivante par un polynôme d'interpolation d'Hermite de degré 5.

RÉSOLUTION. Les éléments soulignés de la table de différences sont les données. On construit les autres éléments au moyen de différences divisées ordinaires.

| | | | | | |
|------------|------------------|-------------------|-----------|------------|--|
| <u>1.3</u> | <u>0.6200860</u> | | | | |
| | | <u>-0.5220232</u> | | | |
| <u>1.3</u> | <u>0.6200860</u> | -0.0897427 | | | |
| | | -0.5489460 | 0.0663657 | | |
| <u>1.6</u> | <u>0.4554022</u> | -0.0698330 | 0.0026663 | | |
| | | <u>-0.5698959</u> | 0.0679655 | -0.0027738 | |
| <u>1.6</u> | <u>0.4554022</u> | -0.0290537 | 0.0010020 | | |
| | | -0.5786120 | 0.0685667 | | |
| <u>1.9</u> | <u>0.2818186</u> | -0.0084837 | | | |
| | | <u>-0.5811571</u> | | | |
| <u>1.9</u> | <u>0.2818186</u> | | | | |

Le choix des éléments le long de la 1ère diagonale descendante produit le polynôme

$$\begin{aligned}
 p(1.5) &= 0.6200860 + (1.5 - 1.3)(-0.5220232) + (1.5 - 1.3)^2(-0.0897427) \\
 &\quad + (1.5 - 1.3)^2(1.5 - 1.6)(0.0663657) + (1.5 - 1.3)^2(1.5 - 1.6)^2(0.0026663) \\
 &\quad + (1.5 - 1.3)^2(1.5 - 1.6)^2(1.5 - 1.9)(-0.0027738) \\
 &= 0.5118277. \quad \square
 \end{aligned}$$

8.6. Interpolation par splines du 3ème degré

On interpole une fonction par des polynômes du 3ème degré par morceaux qui forment une courbe continue qui admet 2 dérivées continues. Les polynômes par morceaux évitent les grandes oscillations de polynômes de degré élevé sur un grand intervalle.

DÉFINITION 8.1. Soit $f(x)$ définie sur l'intervalle $[a, b]$ et un ensemble de nœuds

$$a = x_0 < x_1 < \dots < x_n = b.$$

Un spline d'interpolation S de f est un polynôme du 3ème degré par morceau qui satisfait les conditions suivantes :

- (a) La restriction de $S(x)$ au sous-intervalle $[x_j, x_{j+1}]$ pour chaque $j = 0, 1, \dots, n-1$ est un polynôme du 3ème degré noté $S_j(x)$.
- (b) $S(x_j) = f(x_j)$ pour chaque $j = 0, 1, \dots, n$;
- (c) $S_{j+1}(x_{j+1}) = S_j(x_{j+1})$ pour chaque $j = 0, 1, \dots, n-2$;
- (d) $S'_{j+1}(x_{j+1}) = S'_j(x_{j+1})$ pour chaque $j = 0, 1, \dots, n-2$;
- (e) $S''_{j+1}(x_{j+1}) = S''_j(x_{j+1})$ pour chaque $j = 0, 1, \dots, n-2$;
- (f) $S(x)$ satisfait à l'un des ensembles de conditions aux limites :
 - (i) $S''(x_0) = S''(x_n) = 0$ (frontière libre ou naturelle);
 - (ii) $S'(x_0) = f'(x_0)$ et $S'(x_n) = f'(x_n)$ (pentes aux limites fixées).

On peut définir des splines sous d'autres conditions aux limites. On appelle **spline naturel** un spline aux conditions aux limites libres et **spline aux pentes aux limites fixées** un spline aux pentes aux limites fixées.

Pour construire un spline d'interpolation d'une fonction f , on applique les conditions de la définition au polynôme du 3ème degré

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3,$$

pour $j = 0, 1, \dots, n-1$.

Les splines naturels ou aux pentes aux limites fixées admettent les théorèmes d'existence et d'unicité suivants.

THÉORÈME 8.2 (Spline naturel). *Soit f définie en $a = x_0 < x_1 < \dots < x_n = b$. Alors le spline naturel S qui interpole f aux nœuds x_0, x_1, \dots, x_n avec conditions aux limites $S''(a) = 0$ et $S''(b) = 0$ est unique.*

THÉORÈME 8.3 (Spline aux pentes aux limites fixées). *Soit f définie en $a = x_0 < x_1 < \dots < x_n = b$ et différentiable en a et b . Alors le spline S qui interpole f aux nœuds x_0, x_1, \dots, x_n avec conditions aux limites $S'(a) = f'(a)$ et $S'(b) = f'(b)$ est unique.*

Les commandes de Matlab suivantes tracent les points $(n, \sin n)$ du plan pour $n = 0, 1, \dots, 10$ et échantillonnent $\sin x$ plus finement au 0.25ème par un spline :

```
x = 0:10; y = sin(x);
xx = 0:0.25:10;
yy = spline(x,y,xx);
subplot(2,2,1); plot(x,y,'o',xx,yy);
```

La figure 8.1 montre le résultat..

Les commandes de Matlab suivantes illustrent l'interpolation de données par splines aux pentes nulles aux bouts :

```
x = -4:4; y = [0 .15 1.12 2.36 2.36 1.46 .49 .06 0];
cs = spline(x,[0 y 0]);
xx = linspace(-4,4,101);
plot(x,y,'o',xx,ppval(cs,xx),'-');
```

La figure 8.2 montre le résultat.

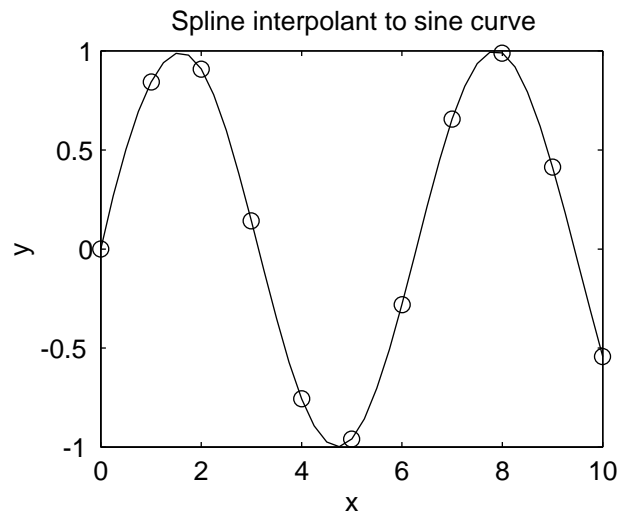
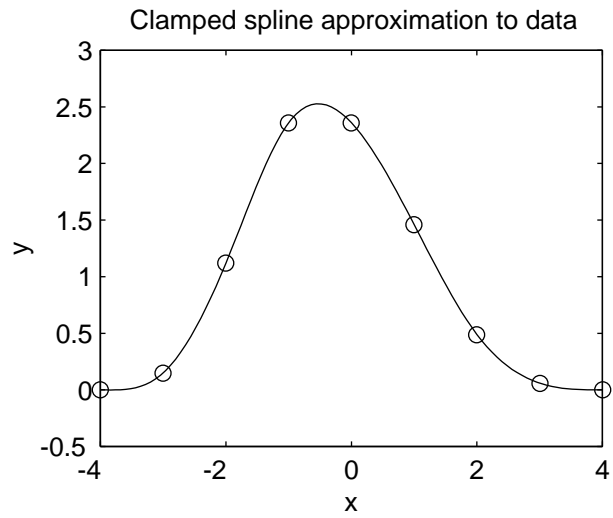
FIGURE 8.1. Interpolation spline de $\sin x$.

FIGURE 8.2. Approximation par splines aux pentes nulles aux bouts.

Différentiation et intégration numériques

9.1. Différentiation numérique

9.1.1. Formule à 2 points pour $f'(x)$. Le polynôme d'interpolation de Lagrange de degré 1 pour $f(x)$ en x_0 et en $x_1 = x_0 + h$ est

$$f(x) = f(x_0) \frac{x - x_1}{-h} + f(x_1) \frac{x - x_0}{h} + \frac{(x - x_0)(x - x_1)}{2!} f''(\xi(x)), \quad x_0 < \xi(x) < x_0 + h.$$

On dérive $f(x)$,

$$f'(x) = f(x_0) \frac{1}{-h} + f(x_1) \frac{1}{h} + \frac{(x - x_1) + (x - x_0)}{2!} f''(\xi(x)) + \frac{(x - x_0)(x - x_1)}{2!} \frac{d}{dx} [f''(\xi(x))]$$

et l'on évalue $f'(x)$ en $x = x_0$. On obtient alors la formule de différentiation numérique à 2 points du premier ordre :

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} - \frac{h}{2} f''(\xi). \quad (9.1)$$

La formule est dite prograde si $h > 0$, et rétrograde si $h < 0$.

9.1.2. Formule à 3 points pour $f'(x)$. Le polynôme d'interpolation de Lagrange de degré 2 pour $f(x)$ en x_0 , $x_1 = x_0 + h$ et $x_2 = x_0 + 2h$ est

$$f(x) = f(x_0) \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + f(x_1) \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + f(x_2) \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} + \frac{(x - x_0)(x - x_1)(x - x_2)}{3!} f'''(\xi(x)),$$

où $x_0 < \xi(x) < x_2$. On dérive $f(x)$ et l'on évalue $f'(x)$ en $x = x_j$:

$$f'(x_j) = f(x_0) \frac{2x_j - x_1 - x_2}{(x_0 - x_1)(x_0 - x_2)} + f(x_1) \frac{2x_j - x_0 - x_2}{(x_1 - x_0)(x_1 - x_2)} + f(x_2) \frac{2x_j - x_0 - x_1}{(x_2 - x_0)(x_2 - x_1)} + \frac{1}{6} f'''(\xi(x_j)) \prod_{k=0, k \neq j}^2 (x_j - x_k).$$

Si $j = 0, 1, 2$, on a 3 formules à 3 points du 2ème ordre :

$$f'(x_0) = f(x_0) \frac{-3h}{2h^2} + f(x_1) \frac{-2h}{-h^2} + f(x_2) \frac{-h}{2h^2} + \frac{2h^2}{6} f'''(\xi_0) = \frac{1}{h} \left[-\frac{3}{2} f(x_0) + 2f(x_1) - \frac{1}{2} f(x_2) \right] + \frac{h^2}{3} f'''(\xi_0),$$

$$\begin{aligned} f'(x_1) &= f(x_0) \frac{-h}{2h^2} + f(x_1) \frac{0}{-h^2} + f(x_2) \frac{h}{2h^2} - \frac{h^2}{6} f'''(\xi_1) \\ &= \frac{1}{h} \left[-\frac{1}{2} f(x_0) + \frac{1}{2} f(x_2) \right] - \frac{h^2}{6} f'''(\xi_1), \end{aligned}$$

et

$$f'(x_2) = \frac{1}{h} \left[\frac{1}{2} f(x_0) - 2f(x_1) + \frac{3}{2} f(x_2) \right] + \frac{h^2}{3} f'''(\xi_2).$$

On récrit ces 3 formules sous forme ordinaire en x_0 :

$$f'(x_0) = \frac{1}{2h} [-3f(x_0) + 4f(x_0 + h) - f(x_0 + 2h)] + \frac{h^2}{3} f'''(\xi_0), \quad (9.2)$$

$$f'(x_0) = \frac{1}{2h} [f(x_0 + h) - f(x_0 - h)] - \frac{h^2}{6} f'''(\xi_1). \quad (9.3)$$

On obtient la 3ème formule en remplaçant h par $-h$ dans (9.2). On remarque que la formule centrée (9.3) est plus précise que (9.2) car son coefficient de l'erreur est la moitié du coefficient d'erreur de (9.2).

9.1.3. Formule à 3 points centrée pour $f''(x)$. On dérive la formule centrée à 3 points pour $f''(x)$ au moyen des développements de Taylor de $f(x+h)$ et de $f(x-h)$:

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h^2 + \frac{1}{6}f'''(x_0)h^3 + \frac{1}{24}f^{(4)}(\xi_0)h^4,$$

$$f(x_0 - h) = f(x_0) - f'(x_0)h + \frac{1}{2}f''(x_0)h^2 - \frac{1}{6}f'''(x_0)h^3 + \frac{1}{24}f^{(4)}(\xi_1)h^4.$$

On additionne ces 2 développements :

$$f(x_0 + h) + f(x_0 - h) = 2f(x_0) + f''(x_0)h^2 + \frac{1}{24} [f^{(4)}(\xi_0) + f^{(4)}(\xi_1)] h^4,$$

et l'on résout pour $f''(x_0)$:

$$f''(x_0) = \frac{1}{h^2} [f(x_0 - h) - 2f(x_0) + f(x_0 + h)] - \frac{1}{24} [f^{(4)}(\xi_0) + f^{(4)}(\xi_1)] h^2.$$

Le théorème de la valeur moyenne 7.5 pour les sommes affirme l'existence d'une valeur ξ , $x_0 - h < \xi < x_0 + h$, telle que

$$\frac{1}{2} [f^{(4)}(\xi_0) + f^{(4)}(\xi_1)] = f^{(4)}(\xi).$$

On obtient donc la formule aux différences centrées à 3 termes du 2ème ordre

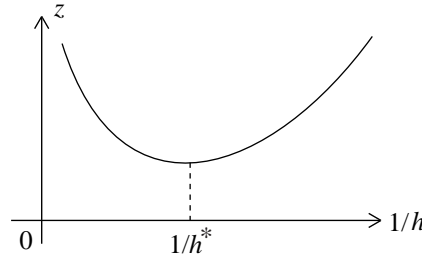
$$f''(x_0) = \frac{1}{h^2} [f(x_0 - h) - 2f(x_0) + f(x_0 + h)] - \frac{h^2}{12} f^{(4)}(\xi). \quad (9.4)$$

9.2. L'effet des erreurs d'arrondi et de méthode

Le pas h qui apparaît au dénominateur des formules de différentiation numérique peut produire de grandes erreurs à cause des erreurs d'arrondi. On considère le cas de la formule centrée à 2 points (9.3) pour $f'(x)$. Les autres formules se traitent de la même façon.

On suppose que l'erreur d'arrondi dans la valeur approchée $\tilde{f}(x_j)$ de $f(x_j)$ est $e(x_j)$. Alors,

$$f(x_0 + h) = \tilde{f}(x_0 + h) + e(x_0 + h), \quad f(x_0 - h) = \tilde{f}(x_0 - h) + e(x_0 - h).$$

FIGURE 9.1. Erreurs de méthode et d'arrondi en fonction de $1/h$.

On substitue ces valeurs dans (9.3), ce qui donne l'erreur totale qui est la somme des erreurs d'arrondi et de méthode :

$$f'(x_0) - \frac{\tilde{f}(x_0 + h) - \tilde{f}(x_0 - h)}{2h} = \frac{e(x_0 + h) - e(x_0 - h)}{2h} - \frac{h^2}{6} f^{(3)}(\xi).$$

On prend le module du second membre et par l'inégalité du triangle on obtient

$$\left| \frac{e(x_0 + h) - e(x_0 - h)}{2h} - \frac{h^2}{6} f^{(3)}(\xi) \right| \leq \frac{1}{2h} (|e(x_0 + h)| + |e(x_0 - h)|) + \frac{h^2}{6} |f^{(3)}(\xi)|.$$

Si

$$|e(x_0 \pm h)| \leq \varepsilon, \quad |f^{(3)}(x)| \leq M,$$

alors

$$\left| f'(x_0) - \frac{\tilde{f}(x_0 + h) - \tilde{f}(x_0 - h)}{2h} \right| \leq \frac{\varepsilon}{h} + \frac{h^2}{6} M.$$

On remarque que l'expression

$$z(h) = \frac{\varepsilon}{h} + \frac{h^2}{6} M$$

commence par décroître puis croît quand $1/h$ croît comme le montre la figure 9.1. Le terme $Mh^2/6$ vient de l'erreur de méthode et le terme ε/h vient de l'erreur d'arrondi.

EXEMPLE 9.1. (a) Soit la fonction $f(x)$ et sa 1ère dérivée $f'(x)$:

$$f(x) = \cos x, \quad f'(x) = -\sin x.$$

Approcher $f'(0.7)$ par la formule à 5 points avec $h = 0.1$ sans le terme de l'erreur de méthode,

$$f'(x) = \frac{1}{12h} [-f(x + 2h) + 8f(x + h) - 8f(x - h) + f(x - 2h)] + \frac{h^4}{30} f^{(5)}(\xi),$$

où ξ dans le terme de l'erreur de méthode satisfait les inégalités $x - 2h \leq \xi \leq x + 2h$.

(b) Si l'erreur d'arrondi dans chaque évaluation de $f(x)$ est bornée par $\varepsilon = 5 \times 10^{-7}$, trouver une borne pour l'erreur totale dans $f'(0.7)$ en additionnant les bornes des erreurs de méthode et d'arrondi.

(c) Trouver la valeur de h qui minimise l'erreur totale.

RÉSOLUTION. (a) Un simple calcul avec la formule donnée, sans l'erreur de méthode, donne l'approximation

$$f'(0.7) \approx -0.644\,215\,542.$$

(b) Puisque la fonction

$$f^{(5)}(x) = -\sin x$$

est négative et décroissante sur l'intervalle $0.5 \leq x \leq 0.9$, alors

$$M = \max_{0.5 \leq x \leq 0.9} |-\sin x| = \sin 0.9 = 0.7833.$$

On a donc la borne de l'erreur totale

$$\begin{aligned} \text{Erreur totale} &\leq \frac{1}{12 \times 0.1} (1 + 8 + 8 + 1) \times 5 \times 10^{-7} + \frac{(0.1)^4}{30} \times 0.7833 \\ &= 7.5000 \times 10^{-6} + 2.6111 \times 10^{-6} \\ &= 1.0111 \times 10^{-5}. \end{aligned}$$

(c) Le minimum de l'erreur totale, comme fonction de h ,

$$\frac{90 \times 10^{-7}}{12h} + \frac{0.7833}{30} h^4,$$

est atteinte en un zéro de sa dérivée par rapport à h ,

$$\frac{d}{dh} \left(\frac{90 \times 10^{-7}}{12h} + \frac{0.7833}{30} h^4 \right) = 0.$$

On fait la dérivée puis l'on multiplie les deux membres par h^2 , ce qui produit une équation du 5ème degré en h :

$$-7.5 \times 10^{-7} + \frac{4 \times 0.7833}{30} h^5 = 0.$$

Donc,

$$\begin{aligned} h &= \left(\frac{7.5 \times 10^{-7} \times 30}{4 \times 0.7833} \right)^{1/5} \\ &= 0.0936 \end{aligned}$$

minimise l'erreur totale. □

9.3. Extrapolation de Richardson

On suppose qu'une formule numérique $N(h)$ approche la valeur exacte M à une erreur près de la forme d'une série en h^j :

$$M = N(h) + K_1 h + K_2 h^2 + K_3 h^3 + \dots,$$

où les constantes K_j sont indépendantes de h . Si l'on calcule $N(h/2)$, on obtient

$$M = N\left(\frac{h}{2}\right) + \frac{1}{2} K_1 h + \frac{1}{4} K_2 h^2 + \frac{1}{8} K_3 h^3 + \dots$$

En soustrayant la 1ère expression de 2 fois la 2ème on élimine l'erreur en h :

$$M = N\left(\frac{h}{2}\right) + \left[N\left(\frac{h}{2}\right) - N(h) \right] + \left(\frac{h^2}{2} - h^2 \right) K_2 + \left(\frac{h^3}{4} - h^3 \right) K_3 + \dots$$

Si l'on pose

$$\begin{aligned} N_1(h) &= N(h), \\ N_2(h) &= N_1\left(\frac{h}{2}\right) + \left[N_1\left(\frac{h}{2}\right) - N_1(h) \right], \end{aligned}$$

la dernière expression pour M devient

$$M = N_2(h) - \frac{1}{2} K_2 h^2 - \frac{3}{4} K_3 h^3 - \dots$$

Maintenant avec $h/4$ on a

$$M = N_2\left(\frac{h}{2}\right) - \frac{1}{8} K_2 h^2 - \frac{3}{32} K_3 h^3 + \dots$$

En soustrayant la 2ème expression de 4 fois la dernière on élimine l'erreur en h^2 et après une division des 2 membres par 3 on obtient :

$$M = \left[N_2\left(\frac{h}{2}\right) + \frac{N_2(h/2) - N_2(h)}{3} \right] + \frac{1}{8} K_3 h^3 + \dots$$

Si l'on pose

$$N_3(h) = N_2\left(\frac{h}{2}\right) + \frac{N_2(h/2) - N_2(h)}{3},$$

on a

$$M = N_3(h) + \frac{1}{8} K_3 h^3 + \dots$$

La présence du nombre $2^{j-1}-1$ au dénominateur du 2ème membre de $N_j(h)$ assure la convergence. On voit comment continuer ce processus appelé **extrapolation de Richardson**.

Un cas important de l'extrapolation de Richardson est la formule de la différence centrée (9.3) pour $f'(x)$:

$$f'(x_0) = \frac{1}{2h} [f(x_0+h) - f(x_0-h)] - \frac{h^2}{6} f'''(x_0) - \frac{h^4}{120} f^{(5)}(x_0) - \dots$$

Puisque le terme de l'erreur ne contient que des puissances paires de h , la convergence de l'extrapolation de Richardson est très rapide. Si l'on pose

$$N_1(h) = N(h) = \frac{1}{2h} [f(x_0+h) - f(x_0-h)],$$

la formule pour $f'(x_0)$ devient

$$f'(x_0) = N_1(h) - \frac{h^2}{6} f'''(x_0) - \frac{h^4}{120} f^{(5)}(x_0) - \dots$$

Si l'on remplace h par $h/2$ dans cette formule, on obtient

$$f'(x_0) = N_1\left(\frac{h}{2}\right) - \frac{h^2}{24} f'''(x_0) - \frac{h^4}{1920} f^{(5)}(x_0) - \dots$$

En soustrayant la 2ème formule pour $f'(x_0)$ de 4 fois la dernière et en divisant le résultat par 3 on a

$$f'(x_0) = N_2(h) - \frac{h^4}{480} f^{(5)}(x_0) + \dots,$$

où

$$N_2(h) = N_1\left(\frac{h}{2}\right) + \frac{N_1(h/2) - N_1(h)}{3}.$$

TABLE 9.1. Extrapolation de Richardson sur la dérivée de $x e^x$.

| | | |
|---------------------------|--------------------------|--------------------------|
| $N_1(0.2) = 22.414\,160$ | $N_2(0.2) = 22.166\,995$ | |
| $N_1(0.1) = 22.228\,786$ | $N_2(0.1) = 22.167\,157$ | $N_3(0.2) = 22.167\,168$ |
| $N_1(0.05) = 22.182\,564$ | | |

La présence du nombre $4^{j-1} - 1$ au dénominateur du 2ème membre de $N_j(h)$ donne une convergence rapide.

EXEMPLE 9.2. Soit

$$f(x) = x e^x.$$

Appliquer l'extrapolation de Richardson pour améliorer la valeur de $f'(x)$ en $x_0 = 2$ obtenue par la formule de la différence centrée (9.3) avec $h = 0.2$.

RÉSOLUTION. On a

$$N_1(0.2) = N(0.2) = \frac{1}{0.4} [f(2.2) - f(1.8)] = 22.414\,160,$$

$$N_1(0.1) = N(0.1) = \frac{1}{0.2} [f(2.1) - f(1.9)] = 22.228\,786,$$

$$N_1(0.05) = N(0.05) = \frac{1}{0.1} [f(2.05) - f(1.95)] = 22.182\,564.$$

Ensuite,

$$N_2(0.2) = N_1(0.1) + \frac{N_1(0.1) - N_1(0.2)}{3} = 22.166\,995,$$

$$N_2(0.1) = N_1(0.05) + \frac{N_1(0.05) - N_1(0.1)}{3} = 22.167\,157.$$

Enfin,

$$N_3(0.2) = N_2(0.1) + \frac{N_2(0.1) - N_2(0.2)}{15} = 22.167\,168,$$

qui est exacte à 6 décimales près. Les résultats sont dans le tableau 9.1. On voit la convergence rapide de l'extrapolation de Richardson sur la formule de la différence centrée. \square

9.4. Intégrations numériques élémentaires

Pour approcher la valeur de l'intégrale définie

$$\int_a^b f(x) dx$$

de la fonction lisse $f(x)$ sur $[a, b]$ où $a < b$, on subdivise cet intervalle en n sous-intervalles de longueur égale $h = (b - a)/n$. On approxime $f(x)$ sur chaque sous-intervalle par un polynôme d'interpolation et l'on intègre tous ces polynômes.

Pour la **méthode du point milieu**, on interpole $f(x)$ sur chaque sous-intervalle $[x_{i-1}, x_i]$ par $f([x_{i-1} + x_i]/2)$, et l'on estime l'intégrale de $f(x)$ sur un sous-intervalle par l'aire d'un rectangle (voir la figure 9.2).

Pour la **méthode du trapèze**, on interpole $f(x)$ sur chaque sous-intervalle $[x_{i-1}, x_i]$ par un polynôme de degré 1 et l'on estime l'intégrale de $f(x)$ sur un sous-intervalle par l'aire d'un trapèze (voir la figure 9.3).

Pour la **méthode de Simpson**, on interpole $f(x)$ sur chaque paire de sous-intervalles $[x_{2i}, x_{2i+1}]$ et $[x_{2i+1}, x_{2i+2}]$ par un polynôme de degré 2 (une parabole) et l'on estime l'intégrale de $f(x)$ sur un sous-intervalle par l'aire sous la parabole (voir la figure 9.4).

9.4.1. Méthode du point milieu. La méthode du point milieu,

$$\int_{x_0}^{x_1} f(x) dx = hf(x_1^*) + \frac{1}{24} f''(\xi)h^3, \quad x_0 < \xi < x_1, \quad (9.5)$$

approche l'intégrale de $f(x)$ sur l'intervalle $x_0 \leq x \leq x_1$ par l'aire du rectangle de hauteur $f(x_1^*)$ et de base $h = x_1 - x_0$, où x_1^* est le point milieu de l'intervalle $[x_0, x_1]$,

$$x_1^* = \frac{x_0 + x_1}{2},$$

(voir la figure 9.2).

Pour dériver la formule (9.5), on développe $f(x)$ en une série de Taylor tronquée de centre $x = x_1^*$,

$$f(x) = f(x_1^*) + f'(x_1^*)(x - x_1^*) + \frac{1}{2} f''(\xi)(x - x_1^*)^2, \quad x_0 < \xi < x_1.$$

L'intégrale de cette expression de x_0 à x_1 donne

$$\begin{aligned} \int_{x_0}^{x_1} f(x) dx &= hf(x_1^*) + \int_{x_0}^{x_1} f'(x_1^*)(x - x_1^*) dx + \frac{1}{2} \int_{x_0}^{x_1} f''(\xi)(x - x_1^*)^2 dx \\ &= hf(x_1^*) + \frac{1}{2} f''(\xi) \int_{x_0}^{x_1} (x - x_1^*)^2 dx. \end{aligned}$$

où l'intégrale du terme linéaire $(x - x_1^*)$ est nulle parce que ce terme est une fonction impaire par rapport au point milieu $x = x_1^*$. On a employé le théorème de la moyenne 7.4 pour intégrales pour intégrer le terme du 2ème degré $(x - x_1^*)^2$ qui ne change pas de signe sur l'intervalle $[x_0, x_1]$. Le résultat suit de la valeur de l'intégrale

$$\frac{1}{2} \int_{x_0}^{x_1} (x - x_1^*)^2 dx = \frac{1}{6} [(x - x_1^*)^3]_{x_0}^{x_1} = \frac{1}{24} h^3.$$

9.4.2. Méthode du trapèze. La méthode du trapèze,

$$\int_{x_0}^{x_1} f(x) dx = \frac{h}{2} [f(x_0) + f(x_1)] - \frac{1}{12} f''(\xi)h^3, \quad x_0 < \xi < x_1, \quad (9.6)$$

approche l'intégrale de $f(x)$ sur l'intervalle $x_0 \leq x \leq x_1$ par l'aire du trapèze de hauteurs $f(x_0)$ et $f(x_1)$ et de base $h = x_1 - x_0$ (voir la figure 9.3).

Pour dériver la formule (9.6), on interpole $f(x)$ en $x = x_0$ et $x = x_1$ par le polynôme de Lagrange linéaire

$$p_1(x) = f(x_0) \frac{x - x_1}{x_0 - x_1} + f(x_1) \frac{x - x_0}{x_1 - x_0}.$$

Donc,

$$f(x) = p_1(x) + \frac{f''(\xi)}{2} (x - x_0)(x - x_1), \quad x_0 < \xi < x_1.$$

Puisque

$$\int_{x_0}^{x_1} p_1(x) dx = \frac{h}{2} [f(x_0) + f(x_1)],$$

on obtient

$$\begin{aligned} & \int_{x_0}^{x_1} f(x) dx - \frac{h}{2} [f(x_0) + f(x_1)] \\ &= \int_{x_0}^{x_1} [f(x) - p_1(x)] dx \\ &= \int_{x_0}^{x_1} \frac{f''(\xi(x))}{2} (x - x_0)(x - x_1) dx \\ & \quad \text{et par le théorème de la moyenne 7.4 pour intégrales} \\ & \quad \text{puisque } (x - x_0)(x - x_1) \leq 0 \text{ sur } [x_0, x_1] \\ &= \frac{f''(\xi)}{2} \int_{x_0}^{x_1} (x - x_0)(x - x_1) dx \\ & \quad \text{et avec } x - x_0 = s, dx = ds, x - x_1 = (x - x_0) - (x_1 - x_0) = s - h \\ &= \frac{f''(\xi)}{2} \int_0^h s(s - h) ds \\ &= \frac{f''(\xi)}{2} \left[\frac{s^3}{3} - \frac{h}{2} s^2 \right]_0^h \\ &= -\frac{f''(\xi)}{12} h^3. \end{aligned}$$

9.4.3. Méthode de Simpson.

La méthode de Simpson

$$\int_{x_0}^{x_2} f(x) dx = \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] - \frac{h^5}{90} f^{(4)}(\xi), \quad x_0 < \xi < x_2, \quad (9.7)$$

approche l'intégrale de $f(x)$ sur l'intervalle $x_0 \leq x \leq x_2$ par l'aire de la parabole qui interpole $f(x)$ en $x = x_0, x_1$ et x_2 (voir la figure 9.4).

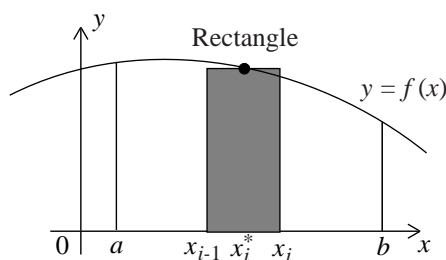
Pour dériver la formule(9.7), on développe $f(x)$ en une série de Taylor tronquée de centre $x = x_1$,

$$f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(x_1)}{2} (x - x_1)^2 + \frac{f'''(x_1)}{6} (x - x_1)^3 + \frac{f^{(4)}(\xi(x))}{24} (x - x_1)^4.$$

On intègre cette expression de x_0 à x_2 en notant que les termes $(x - x_1)$ et $(x - x_1)^3$ sont des fonctions impaires par rapport au point $x = x_1$ de sorte que leurs intégrales sont nulles. On obtient donc

$$\begin{aligned} \int_{x_0}^{x_2} f(x) dx &= \left[f(x_1)x + \frac{f''(x_1)}{6} (x - x_1)^3 + \frac{f^{(4)}(\xi_1)}{120} (x - x_1)^5 \right]_{x_0}^{x_2} \\ &= 2hf(x_1) + \frac{h^3}{3} f''(x_1) + \frac{f^{(4)}(\xi_1)}{60} h^5, \end{aligned}$$

où l'on a employé le théorème de la moyenne 7.4 pour intégrales pour intégrer le terme de l'erreur puisque le facteur $(x - x_1)^4$ ne change pas de signe sur l'intervalle $[x_0, x_2]$.

FIGURE 9.2. Le i ème panneau de la méthode du point milieu.

En substituant la formule de différence centrée à 2 points (9.4) pour $f''(x_1)$ avec $f(x_0)$, $f(x_1)$ et $f(x_2)$:

$$f''(x_1) = \frac{1}{h^2} [f(x_0) - 2f(x_1) + f(x_2)] - \frac{1}{12} f^{(4)}(\xi_2) h^2,$$

on obtient

$$\int_{x_0}^{x_2} f(x) dx = \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] - \frac{h^5}{12} \left[\frac{1}{3} f^{(4)}(\xi_2) - \frac{1}{5} f^{(4)}(\xi_1) \right].$$

Ici on ne peut pas employer le théorème de la moyenne 7.4 pour sommes pour exprimer le terme de l'erreur sous la forme $f^{(4)}(\xi)$ évalué en un seul point puisque les poids $1/3$ et $-1/5$ ne sont pas de même signe. Cependant, puisque la formule est exacte pour les polynômes de degré au plus 4, pour obtenir le facteur $1/90$ il suffit d'appliquer la formule sur le monôme $f(x) = x^4$ et, par simplicité, d'intégrer de $-h$ à h :

$$\begin{aligned} \int_{-h}^h x^4 dx &= \frac{h}{3} [(-h)^4 + 4(0)^4 + h^4] + k f^{(4)}(\xi) \\ &= \frac{2}{3} h^5 + 4!k = \frac{2}{5} h^5, \end{aligned}$$

où le dernier terme est la valeur exacte de l'intégrale. Il suit que

$$k = \frac{1}{4!} \left[\frac{2}{5} - \frac{2}{3} \right] h^5 = -\frac{1}{90} h^5,$$

implique (9.7).

9.5. La méthode des points milieux (composée)

On divise l'intervalle $[a, b]$ en n sous-intervalles de longueur $h = (b - a)/n$:

$$x_0 = a, \quad x_1 = a + h, \quad \dots, \quad x_i = a + ih, \quad \dots, \quad x_n = b.$$

On approche l'intégrale de $f(x)$ sur le sous-intervalle $[x_{i-1}, x_i]$ par l'aire signée du rectangle de base $[x_{i-1}, x_i]$ et de hauteur $f(x_i^*)$ où

$$x_i^* = \frac{1}{2} (x_{i-1} + x_i)$$

est le point milieu du segment $[x_{i-1}, x_i]$ tel que le montre la figure 9.2. Alors, par la méthode du point milieu (9.5) sur l'intervalle $[x_{i-1}, x_i]$ on a

$$\int_{x_{i-1}}^{x_i} f(x) dx = h f(x_i^*) + \frac{1}{24} f''(\xi_i) h^3, \quad x_{i-1} < \xi_i < x_i.$$

On somme cette formule sur tous les sous-intervalles :

$$\int_a^b f(x) dx = h \sum_{i=1}^n f(x_i^*) + \frac{h^3}{24} \sum_{i=1}^n f''(\xi_i).$$

On multiplie et divise le terme de l'erreur par n , on lui applique le théorème de la moyenne 7.5 pour les sommes et l'on emploie la relation $nh = b - a$ pour obtenir

$$\frac{nh^3}{24} \sum_{i=1}^n \frac{1}{n} f''(\xi_i) = \frac{(b-a)h^2}{24} f''(\xi), \quad a < \xi < b.$$

On a donc dérivé la **méthode des points milieux** (composée) :

$$\int_a^b f(x) dx = h[f(x_1^*) + f(x_2^*) + \cdots + f(x_n^*)] + \frac{(b-a)h^2}{24} f''(\xi), \quad a < \xi < b. \quad (9.8)$$

On voit que la méthode des points milieux est d'ordre $O(h^2)$, donc exacte pour les polynômes de degré 1.

EXEMPLE 9.3. Approcher l'intégrale

$$I = \int_0^1 e^{x^2} dx$$

par la méthode des points milieux avec un pas h tel que le module de l'erreur de méthode soit borné par 10^{-4} .

RÉSOLUTION. Puisque

$$f(x) = e^{x^2} \quad \text{et} \quad f''(x) = (2 + 4x^2) e^{x^2},$$

alors

$$0 \leq f''(x) \leq 6e \quad \text{pour} \quad x \in [0, 1].$$

Donc, le module de l'erreur de méthode satisfait l'inégalité

$$|\epsilon_M| \leq \frac{1}{24} 6e(1-0)h^2 = \frac{1}{4} eh^2 < 10^{-4}.$$

Il suit que

$$h < 0.0121, \quad \frac{1}{h} = 82.4361.$$

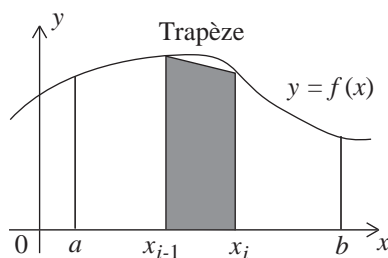
On prend $n = 83 \geq 1/h = 82.4361$ et $h = 1/83$. La valeur approchée de I est

$$\begin{aligned} I &\approx \frac{1}{83} \left[e^{(0.5/83)^2} + e^{(1.5/83)^2} + \cdots + e^{(13590.5/83)^2} + e^{(82.5/83)^2} \right] \\ &\approx 1.46262 \quad \square \end{aligned}$$

Les commandes de Matlab font l'intégration par la méthode des points milieux.

```
x = 0.5:82.5; y = exp((x/83).^2);
z = 1/83*sum(y)
z = 1.4626
```

□

FIGURE 9.3. Le i ème panneau de la méthode des trapèzes.

9.6. La méthode des trapèzes (composée)

On divise l'intervalle $[a, b]$ en n sous-intervalles de longueur $h = (b - a)/n$:

$$x_0 = a, \quad x_1 = a + h, \quad \dots, \quad x_i = a + ih, \quad \dots, \quad x_n = b.$$

On approche l'intégrale de $f(x)$ sur le sous-intervalle $[x_{i-1}, x_i]$ par l'aire signée du trapèze aux sommets

$$(x_{i-1}, 0), \quad (x_i, 0), \quad (x_i, f(x_i)), \quad (x_{i-1}, f(x_{i-1})),$$

comme à la figure 9.3. Par la méthode du trapèze (9.6) on a

$$\int_{x_{i-1}}^{x_i} f(x) dx = \frac{h}{2} [f(x_{i-1}) + f(x_i)] - \frac{h^3}{12} f''(\xi_i).$$

On somme cette formule sur tous les sous-intervalles :

$$\int_a^b f(x) dx = \frac{h}{2} \sum_{i=1}^n [f(x_{i-1}) + f(x_i)] - \frac{h^3}{12} \sum_{i=1}^n f''(\xi_i).$$

On multiplie et divise le terme de l'erreur par n , on lui applique le théorème de la moyenne 7.5 pour les sommes et l'on emploie la relation $nh = b - a$ pour obtenir

$$-\frac{nh^3}{12} \sum_{i=1}^n \frac{1}{n} f''(\xi_i) = -\frac{(b-a)h^2}{12} f''(\xi), \quad a < \xi < b.$$

On a donc dérivé la **méthode des trapèzes** (composée) :

$$\begin{aligned} \int_a^b f(x) dx &= \frac{h}{2} [f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-2}) \\ &\quad + 2f(x_{n-1}) + f(x_n)] - \frac{(b-a)h^2}{12} f''(\xi), \quad a < \xi < b. \end{aligned} \quad (9.9)$$

On voit que la méthode des trapèzes est d'ordre $O(h^2)$, donc exacte pour les polynômes de degré 1. Le module de son erreur de méthode est le double du module correspondant de la méthode des points milieux.

EXEMPLE 9.4. Approcher l'intégrale

$$I = \int_0^1 e^{x^2} dx$$

par la méthode des trapèzes avec un pas h tel que le module de l'erreur de méthode soit borné par 10^{-4} . Comparer vos résultats avec ceux des exemples 9.3 et 9.6.

RÉSOLUTION. Puisque

$$f(x) = e^{x^2} \quad \text{and} \quad f''(x) = (2 + 4x^2)e^{x^2},$$

alors

$$0 \leq f''(x) \leq 6e \quad \text{for} \quad x \in [0, 1].$$

Donc,

$$|\epsilon_T| \leq \frac{1}{12} 6e(1-0)h^2 = \frac{1}{2}eh^2 < 10^{-4}, \quad \text{c'est-à-dire,} \quad h < 0.008577638.$$

On prend $n = 117 \geq 1/h = 116.6$ (83 suffisaient pour la méthode des points milieux). La valeur approchée de I est

$$\begin{aligned} I &\approx \frac{1}{117 \times 2} \left[e^{(0/117)^2} + 2e^{(1/117)^2} + 2e^{(2/117)^2} + \dots \right. \\ &\quad \left. + 2e^{(115/117)^2} + 2e^{(116/117)^2} + e^{(117/117)^2} \right] \\ &= 1.46268. \end{aligned}$$

Les commandes de Matlab suivantes calculent les valeurs y_k par intégration numérique par les trapèzes aux nœuds $k/117$, $k = 0, 1, \dots, 117$, avec le pas $h = 1/117$.

```
x = 0:117; y = exp((x/117).^2);
z = trapz(x,y)/117
z = 1.4627
```

□

EXEMPLE 9.5. Quel est le nombre de sous-intervalles minimum pour approcher l'intégrale

$$I = \int_1^2 \left[x^2 - \frac{1}{12}(x-1.5)^4 \right] dx$$

par la méthode des trapèzes à 10^{-3} près ?

RÉSOLUTION. Notons

$$f(x) = x^2 - \frac{1}{12}(x-1.5)^4.$$

Alors

$$f''(x) = 2 - (x-1.5)^2.$$

On voit que

$$M = \max_{1 \leq x \leq 2} f''(x) = f(1.5) = 2.$$

Pour borner le module de l'erreur de méthode par 10^{-3} , il faut

$$\left| \frac{(b-a)h^2}{12} f''(\xi) \right| \leq \frac{h^2}{12} M = \frac{h^2}{6} \leq 10^{-3},$$

d'où

$$h \leq \sqrt{6 \times 10^{-3}} = 0.0775$$

et

$$\frac{1}{h} = 12.9099 \leq n = 13.$$

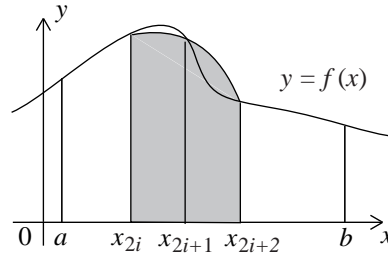


FIGURE 9.4. Un panneau double de la méthode de Simpson.

Il suffit donc de prendre

$$h = \frac{1}{13}, \quad n = 13. \quad \square$$

9.7. La méthode de Simpson (composée)

On divise l'intervalle $[a, b]$ en en un nombre **pair** $n = 2m$ de sous-intervalles de longueur $h = (b - a)/(2m)$:

$$x_0 = a, \quad x_1 = a + h, \quad \dots, \quad x_i = a + i h, \quad \dots, \quad x_{2m} = b.$$

On interpole la fonction $f(x)$ sur le sous-intervalle $[x_{2i}, x_{2i+2}]$ par le polynôme du 2ème degré qui passe par les points

$$(x_{2i}, f(x_{2i})), \quad (x_{2i+1}, f(x_{2i+1})), \quad (x_{2i+2}, f(x_{2i+2})),$$

tel qu'illustré dans la figure 9.4.

Par la méthode de Simpson (9.7) :

$$\int_{x_{2i}}^{x_{2i+2}} f(x) dx = \frac{h}{3} [f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})] - \frac{h^5}{90} f^{(4)}(\xi_i), \quad x_{2i} < \xi < x_{2i+2}.$$

On somme cette formule sur tous les sous-intervalles :

$$\int_a^b f(x) dx = \frac{h}{3} \sum_{i=1}^m [f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})] - \frac{h^5}{90} \sum_{i=1}^m f^{(4)}(\xi_i).$$

On multiplie et divise le terme de l'erreur par m , on lui applique le théorème de la moyenne 7.5 pour les sommes et l'on emploie la relation $2mh = nh = b - a$ pour obtenir

$$-\frac{2mh^5}{2 \times 90} \sum_{i=1}^m \frac{1}{m} f^{(4)}(\xi_i) = -\frac{(b-a)h^4}{180} f^{(4)}(\xi), \quad a < \xi < b.$$

On a donc dérivé la **méthode de Simpson** (composée) :

$$\begin{aligned} \int_a^b f(x) dx &= \frac{h}{3} [f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \dots \\ &+ 2f(x_{2m-2}) + 4f(x_{2m-1}) + f(x_{2m})] - \frac{(b-a)h^4}{180} f^{(4)}(\xi), \quad a < \xi < b. \end{aligned} \quad (9.10)$$

On voit que la méthode de Simpson est d'ordre $O(h^4)$, donc exacte pour les polynômes de degré au plus 3.

EXEMPLE 9.6. Approcher l'intégrale

$$I = \int_0^1 e^{x^2} dx$$

par la méthode de Simpson (composée) au 10^{-4} près. Comparer vos résultats avec ceux des exemples 9.3 et 9.4.

RÉSOLUTION. De

$$f(x) = e^{x^2} \quad \text{et} \quad f^{(4)}(x) = 4e^{x^2} (3 + 12x^2 + 4x^4)$$

il suit que

$$0 \leq f^{(4)}(x) \leq 76e \quad \text{sur} \quad [0, 1].$$

Donc le module de l'erreur de méthode est au plus $\frac{76}{180} e(1-0)h^4$. Par conséquent, h doit satisfaire l'inégalité

$$\frac{76}{180} eh^4 < 10^{-4}, \quad \text{c'est-à-dire,} \quad h < 0.096\,614\,232.$$

Pour satisfaire l'inégalité

$$2m \geq \frac{1}{h} = 10.4$$

on prend

$$n = 2m = 12 \quad \text{et} \quad h = \frac{1}{12}.$$

L'approximation est

$$\begin{aligned} I &\approx \frac{1}{12 \times 3} \left[e^{(0/12)^2} + 4e^{(1/12)^2} + 2e^{(2/12)^2} + \dots + 2e^{(10/12)^2} + 4e^{(11/12)^2} + e^{(12/12)^2} \right] \\ &= 1.46267. \end{aligned}$$

On a obtenu une valeur semblable aux valeurs obtenues aux exemples 9.3 et 9.4. Cependant, le nombre d'opérations arithmétiques est très inférieur avec la méthode de Simpson (ce qui réduit le coût et les erreurs d'arrondi). En général, en général, on préfère la méthode de Simpson aux méthodes des points milieux et des trapèzes. \square

EXEMPLE 9.7. Approcher l'intégrale

$$I = \int_0^2 \sqrt{1 + \cos^2 x} dx$$

par la méthode de Simpson (composée) au 0.0001 près.

RÉSOLUTION. On détermine le pas h tel que le module de l'erreur de méthode $|\epsilon_S|$ soit borné par 0.0001. La 4ème dérivée de

$$f(x) = \sqrt{1 + \cos^2 x}$$

est

$$\begin{aligned} f^{(4)}(x) &= \frac{-3 \cos^4(x)}{(1 + \cos^2(x))^{3/2}} + \frac{4 \cos^2(x)}{(1 + \cos^2(x))^{1/2}} - \frac{18 \cos^4(x) \sin^2(x)}{(1 + \cos^2(x))^{5/2}} \\ &+ \frac{22 \cos^2(x) \sin^2(x)}{(1 + \cos^2(x))^{3/2}} - \frac{4 \sin^2(x)}{(1 + \cos^2(x))^{1/2}} - \frac{15 \cos^4(x) \sin^4(x)}{(1 + \cos^2(x))^{7/2}} \\ &+ \frac{18 \cos^2(x) \sin^4(x)}{(1 + \cos^2(x))^{5/2}} - \frac{3 \sin^4(x)}{(1 + \cos^2(x))^{3/2}}. \end{aligned}$$

Puisque chaque dénominateur est supérieur à 1, il suit que

$$|f^{(4)}(x)| \leq 3 + 4 + 18 + 22 + 4 + 15 + 18 + 3 = 87.$$

Pour satisfaire l'inégalité

$$|\epsilon_S| < \frac{87}{180} (2-0)h^4.$$

il faut que

$$h < 0.100\,851\,140, \quad \frac{1}{h} > 9.915\,604\,269.$$

Pour avoir $2m \geq 2/h = 2 \times 9.9$ on prend $n = 2m = 20$ et $h = 1/10$. L'approximation est

$$\begin{aligned} I &\approx \frac{1}{20 \times 3} \left[\sqrt{1 + \cos^2(0)} + 4 \sqrt{1 + \cos^2(0.1)} + 2 \sqrt{1 + \cos^2(0.2)} + \dots \right. \\ &\quad \left. + 2 \sqrt{1 + \cos^2(1.8)} + 4 \sqrt{1 + \cos^2(1.9)} + \sqrt{1 + \cos^2(2)} \right] \\ &= 2.35169. \quad \square \end{aligned}$$

9.8. Intégration de Romberg pour la méthode des trapèzes

Soit l'intégrale

$$I = \int_a^b f(x) dx$$

et l'approximation $R_{k,1} \approx I$ obtenue par formule des trapèzes avec le pas h_k . L'intégration de Romberg emploie l'extrapolation de Richardson pour améliorer $R_{k,1}$. On peut montrer que

$$I = R_{k,1} + K_1 h_k^2 + K_2 h_k^4 + K_3 h_k^6 + \dots,$$

où les constantes K_j sont indépendantes de h_k . Au moyen des pas

$$h_1 = h, \quad h_2 = \frac{h}{2}, \quad h_3 = \frac{h}{2^2}, \quad \dots, \quad h_k = \frac{h}{2^{k-1}}, \quad \dots,$$

on peut annuler les erreurs d'ordre h^2 , h^4 , etc. de la façon suivante. Si l'on a calculé $R_{k,1}$ et $R_{k+1,1}$, alors

$$I = R_{k,1} + K_1 h_k^2 + K_2 h_k^4 + K_3 h_k^6 + \dots$$

et

$$I = R_{k+1,1} + K_1 \frac{h_k^2}{4} + K_2 \frac{h_k^2}{16} + K_3 \frac{h_k^2}{64} + \dots$$

On soustrait la 1ère expression pour I de 4 fois la 2ème et l'on divise par 3 :

$$I = \left[R_{k+1,1} + \frac{R_{k+1,1} - R_{k,1}}{3} \right] + \frac{K_2}{3} \left[\frac{1}{4} - 1 \right] h_k^4 + \frac{K_3}{3} \left[\frac{1}{16} - 1 \right] h_k^4 + \dots$$

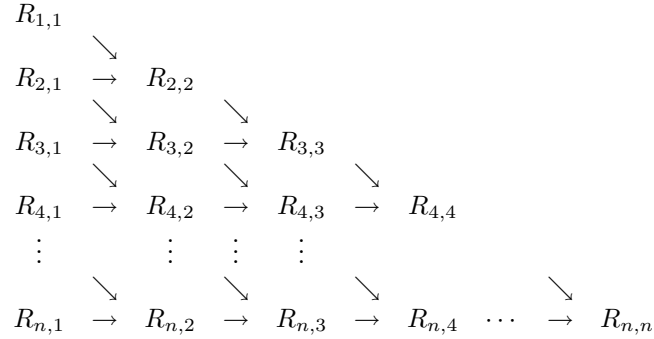
Posons

$$R_{k,2} = R_{k,1} + \frac{R_{k,1} - R_{k-1,1}}{3}$$

et, en général,

$$R_{k,j} = R_{k,j-1} + \frac{R_{k,j-1} - R_{k-1,j-1}}{4^{j-1} - 1}.$$

Alors $R_{k,j}$ est une meilleure approximation de I que $R_{k,j-1}$ et $R_{k-1,j-1}$. On montre les relations entre les $R_{k,j}$ dans le tableau 9.2.

TABLE 9.2. Table d'intégration de Romberg à n niveaux.

EXEMPLE 9.8. Approcher l'intégrale

$$I = \int_0^{\pi/4} \tan x \, dx$$

par une intégration de Romberg à 6 niveaux avec $h_1 = h = \pi/4$.

RÉSOLUTION. Un simple programme de Matlab produit les résultats du tableau.

Romberg integration table:

```

0.39269908
0.35901083  0.34778141
0.34975833  0.34667417  0.34660035
0.34737499  0.34658054  0.34657430  0.34657388
0.34677428  0.34657404  0.34657360  0.34657359  0.34657359
0.34662378  0.34657362  0.34657359  0.34657359  0.34657359  0.34657359

```

□

9.9. Quadratures adaptées

Les méthodes (composées) à pas constant qui sont exactes pour les polynômes de degré d sont efficaces si la $(d+1)$ ème dérivée $f^{(d+1)}$ se comporte uniformément bien sur l'intervalle d'intégration $[a, b]$. Cependant, si le module de cette dérivée admet une grande variation sur cet intervalle, le contrôle de l'erreur surestimer le nombre d'évaluations de la fonction à intégrer. De fait, le nombre n de nœuds est déterminé la borne globale M_{d+1} de $f^{(d+1)}$ sur tout l'intervalle. Là où $|f^{(d+1)}|$ est faible devant M_{d+1} , les sous-intervalles sont (possiblement) beaucoup plus courts qu'il le faut. Les *quadratures adaptées* remédient à ce problème en dépistant les intervalles où $f(x)$ se comporte mal et en réduisant la longueur des sous-intervalles s'il le faut.

On utilise la formule de Simpson à titre d'exemple :

$$I =: \int_a^b f(x) \, dx = S(a, b) - \frac{h^5}{90} f^{(4)}(\xi), \quad 0 < \xi < b,$$

où

$$S(a, b) = \frac{h}{3} [f(a) + 4f(a+h) + f(b)], \quad h = \frac{b-a}{2}.$$

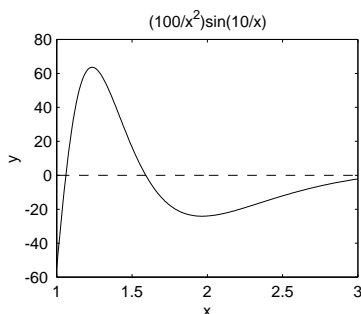


FIGURE 9.5. Une fonction à variation rapide pour une quadrature adaptée.

Une quadrature adaptée utilise un grand h là où $|f^{(4)}(x)|$ est faible et un petit h là où $|f^{(4)}(x)|$ est grand pour minimiser l'erreur uniformément. Un processus simple pour estimer l'erreur consiste à faire les calculs avec h et $h/2$:

$$I = S(a, b) - \frac{h^5}{90} f^{(4)}(\xi_1), \quad (9.11)$$

$$I = S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right) - \frac{2}{32} \frac{h^5}{90} f^{(4)}(\xi_2). \quad (9.12)$$

Sous l'hypothèse que

$$f^{(4)}(\xi_2) \approx f^{(4)}(\xi_1),$$

on soustrait la 2ème expression pour I de la 1ère pour obtenir une expression pour le terme de l'erreur :

$$\frac{h^5}{90} f^{(4)}(\xi_1) \approx \frac{16}{15} \left[S(a, b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right) \right].$$

En insérant cette expression dans (9.12), on peut estimer l'erreur absolue :

$$\left| I - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right) \right| \approx \frac{1}{15} \left| S(a, b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right) \right|.$$

Si le second membre est inférieur à la tolérance désirée, alors

$$S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right)$$

approxime bien la valeur de I .

La quadrature de Simpson adaptée est souvent meilleure que la méthode de Simpson (composée). Par exemple, on intègre la fonction

$$f(x) = \frac{100}{x^2} \sin\left(\frac{10}{x}\right), \quad 1 \leq x \leq 3,$$

tracée dans la figure 9.5. Sous la tolérance 10^{-4} , la quadrature de Simpson adaptée emploie 23 sous-intervalles et requiert 93 évaluations de f . D'autre part, la méthode de Simpson (composée) au pas constant $h = 1/88$ requiert 177 évaluations de f . On voit sur la figure que f varie rapidement dans l'intervalle $[1, 1.5]$. La quadrature adaptée requiert 11 sous-intervalles sur le court intervalle $[1, 1.5]$ et seulement 12 sur l'intervalle plus long $[1.5, 3]$.

Les routines de quadrature MATLAB `quad`, `quadl` et `dblquad` sont des quadratures adaptées.

La méthode de Simpson de Matlab adaptée `quad` et la méthode de Newton-Cotes à 8 panneaux `quad8` évaluent l'intégrale

$$I = \int_0^{\pi/2} \sin x \, dx$$

de la façon suivante :

```
>> v1 = quad('sin',0,pi/2)
v1 = 1.00000829552397
>> v2 = quad8('sin',0,pi/2)
v2 = 1.00000000000000
```

avec une erreur relative 10^{-3} .

9.10. Quadratures gaussiennes

Les quadratures gaussiennes sont les formules d'intégration les plus précises pour un nombre de nœuds donné. La quadrature gaussienne à n points approche la valeur de l'intégrale de $f(x)$ sur l'intervalle $-1 \leq x \leq 1$ par la formule

$$\int_{-1}^1 f(x) \, dx \approx \sum_{i=1}^n w_i f(x_i) \quad (9.13)$$

où les nœuds x_i sont les zéros du polynôme de Legendre $P_n(x)$ de degré n .

La quadrature gaussienne à 2 points est donnée par la formule

$$\int_{-1}^1 f(x) \, dx = f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right).$$

La quadrature gaussienne à 3 points est donnée par la formule

$$\int_{-1}^1 f(x) \, dx = \frac{5}{9}f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9}f(0) + \frac{5}{9}f\left(\sqrt{\frac{3}{5}}\right).$$

Les nœuds x_i , les poids w_i et la précision $2n - 1$ des quadratures gaussiennes à n points pour $n = 1, 2, \dots, 5$ se trouvent dans le tableau 9.3.

L'erreur dans la formule à n points est

$$E_n(f) = \frac{2}{(2n+1)!} \left[\frac{2^n (n!)^2}{(2n)!} \right]^2 f^{(2n)}(\xi), \quad -1 < \xi < 1.$$

La formule est donc exacte pour les polynômes de degré au plus $2n - 1$.

On dérive les quadratures gaussiennes dans la section 6.6 au moyen des relations d'orthogonalité des polynômes de Legendre. On peut aussi obtenir ces quadratures par une intégration des éléments de la base de Lagrange sur $-1 \leq x \leq 1$ pour chaque zéro x_i du polynôme $P_n(x)$ de Legendre :

$$w_i = \int_{-1}^1 \prod_{j=1, j \neq i}^n \frac{x - x_j}{x_i - x_j} \, dx.$$

On trouve des exemples à la section 6.6 et des exercices parmi les exercices pour le chapitre 6.

TABLE 9.3. Nœuds x_i , poids w_i et précision $2n - 1$ des quadratures gaussiennes à n points.

| n | x_i | w_i | Précision $2n - 1$ |
|-----|------------------|-----------------|--------------------|
| 2 | $-1/\sqrt{3}$ | 1 | 3 |
| | $1/\sqrt{3}$ | 1 | |
| 3 | $-\sqrt{3/5}$ | 5/9 | 5 |
| | 0 | 8/9 | |
| | $\sqrt{3/5}$ | 5/9 | |
| 4 | -0.861 136 311 6 | 0.347 854 845 1 | 7 |
| | -0.339 981 043 6 | 0.652 145 154 9 | |
| | 0.339 981 043 6 | 0.652 145 154 9 | |
| | 0.861 136 311 6 | 0.347 854 845 1 | |
| 5 | -0.906 179 845 9 | 0.236 926 885 1 | 9 |
| | -0.538 469 310 1 | 0.478 628 670 5 | |
| | 0 | 0,568 888 888 9 | |
| | 0.538 469 310 1 | 0.478 628 670 5 | |
| | 0.906 179 845 9 | 0.236 926 885 1 | |

Dans les applications, on subdivise l'intervalle d'intégration $[a, b]$ et l'on applique une quadrature gaussienne sur chaque sous-intervalle avec le changement de variable approprié comme à l'exemple 6.11.

Résolution numérique d'équations différentielles

10.1. Problèmes à valeur initiale

On considère le problème différentiel du 1er ordre à valeur initiale :

$$y' = f(x, y), \quad y(x_0) = y_0. \quad (10.1)$$

Pour approcher la solution $y(x)$ de (10.1) sur l'intervalle $a \leq x \leq b$, on choisit $N + 1$ points distincts, x_0, x_1, \dots, x_N , tels que $a = x_0 < x_1 < x_2 < \dots < x_N = b$, et l'on construit les valeurs approchées $y_n \approx y(x_n)$ en x_n , $n = 0, 1, \dots, N$.

Il est important de savoir si une *petite perturbation* de (10.1) va produire une *grande variation* de la solution. Si tel est le cas, il est très peu probable qu'on puisse trouver une bonne approximation de la solution de (10.1). Les erreurs d'arrondi dans le calcul de $f(x, y)$ et l'évaluation des conditions initiales sont des perturbations de (10.1).

DÉFINITION 10.1. Le problème (10.1) est bien posé au sens d'Hadamard s'il admet une et une seule solution et qu'une petite perturbation du problème ne produit qu'un petit changement de la solution.

Le théorème suivant donne des conditions suffisantes pour avoir un problème bien posé.

THÉORÈME 10.1. *Soit*

$$D = \{(x, y) : a \leq x \leq b \text{ et } -\infty < y < \infty\}.$$

Si $f(x, y)$ est continue en D et lipschitzienne par rapport à y :

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2| \quad (10.2)$$

pour tout (x, y_1) et (x, y_2) en D , où L est la constante de Lipschitz, alors le problème à valeur initiale (10.1) est bien posé.

On supposera toujours que les conditions du théorème 10.1 sont valides et que le problème (10.1) est bien posé. De plus on supposera que $f(x, y)$ admet des dérivées partielles d'ordre arbitraire.

On emploiera la notation suivante pour les méthodes numériques pour résoudre (10.1) :

- $h > 0$ est le *pas d'intégration*;
- $x_n = x_0 + nh$ est le *nième nœud*;
- $y(x_n)$ est la *solution exacte* en x_n ;
- y_n est la *solution numérique* en x_n ;
- $f_n = f(x_n, y_n)$ est la valeur numérique de $f(x, y)$ en (x_n, y_n) .

Une fonction $g(x)$ est d'ordre *ordre* p quand $x \rightarrow x_0$, noté $g \in O(|x - x_0|^p)$, si

$$|g(x)| < M|x - x_0|^p, \quad M \text{ une constante,}$$

pour tout x près de x_0 .

10.2. Méthodes explicites à un pas

On commence par la méthode explicite la plus simple.

10.2.1. Méthode d'Euler. Pour approcher la solution $y(x)$ de (10.1) sur l'intervalle $a \leq x \leq b$, on choisit $N + 1$ points distincts, x_0, x_1, \dots, x_N , tels que $a = x_0 < x_1 < x_2 < \dots < x_N = b$, et l'on pose $h = (x_N - x_0)/N$. Par le théorème de Taylor,

$$y(x_{n+1}) = y(x_n) + y'(x_n)(x_{n+1} - x_n) + \frac{y''(\xi_n)}{2}(x_{n+1} - x_n)^2$$

avec ξ_n entre x_n et x_{n+1} pour $n = 0, 1, \dots, N$. Puisque $y'(x_n) = f(x_n, y(x_n))$ et $x_{n+1} - x_n = h$, on a

$$y(x_{n+1}) = y(x_n) + f(x_n, y(x_n))h + \frac{y''(\xi_n)}{2}h^2.$$

On obtient la méthode d'Euler :

$$y_{n+1} = y_n + hf(x_n, y_n), \quad (10.3)$$

si l'on omet le terme

$$\frac{y''(\xi_n)}{2}h^2,$$

d'ordre $O(h^2)$, appelé **erreur de méthode**.

On a l'algorithme de la **méthode d'Euler** :

- (1) Prendre h tel que $N = (x_N - x_0)/h$ soit un entier.
- (2) Débuter avec y_0 et, pour $n = 0, 1, \dots, N$, itérer

$$y_{n+1} = y_n + hf(x_0 + nh, y_n). \quad (10.4)$$

Alors, y_n est une approximation de $y(x_n)$.

EXEMPLE 10.1. Approcher la solution du problème à valeur initiale

$$y'(x) = 0.2xy, \quad y(1) = 1, \quad (10.5)$$

sur l'intervalle $1 \leq x \leq 1.5$ par la méthode d'Euler avec $h = 0.1$.

RÉSOLUTION. On a

$$x_0 = 1, \quad x_N = 1.5, \quad y_0 = 1, \quad f(x, y) = 0.2xy.$$

Alors

$$x_n = x_0 + hn = 1 + 0.1n, \quad N = \frac{1.5 - 1}{0.1} = 5,$$

et

$$y_{n+1} = y_n + 0.1 \times 0.2(1 + 0.1n)y_n, \quad \text{avec } y_0 = 1,$$

pour $n = 0, 1, \dots, 4$. Les résultats numériques sont dans le tableau 10.1. On remarque que l'équation différentielle (10.5) est séparable et que sa solution (unique) est

$$y(x) = e^{(0.1x^2 - 0.1)}.$$

TABLE 10.1. Résultats numériques pour l'exemple 10.1.

| n | x_n | y_n | $y(x_n)$ | Erreur absolue | Erreur relative |
|-----|-------|--------|----------|----------------|-----------------|
| 0 | 1.00 | 1.0000 | 1.0000 | 0.0000 | 0.00 |
| 1 | 1.10 | 1.0200 | 1.0212 | 0.0012 | 0.12 |
| 2 | 1.20 | 1.0424 | 1.0450 | 0.0025 | 0.24 |
| 3 | 1.30 | 1.0675 | 1.0714 | 0.0040 | 0.37 |
| 4 | 1.40 | 1.0952 | 1.1008 | 0.0055 | 0.50 |
| 5 | 1.50 | 1.1259 | 1.1331 | 0.0073 | 0.64 |

TABLE 10.2. Résultats numériques pour l'exemple 10.2.

| n | x_n | y_n | $y(x_n)$ | Erreur absolue | Erreur relative |
|-----|-------|--------|----------|----------------|-----------------|
| 0 | 1.00 | 1.0000 | 1.0000 | 0.0000 | 0.00 |
| 1 | 1.10 | 1.2000 | 1.2337 | 0.0337 | 2.73 |
| 2 | 1.20 | 1.4640 | 1.5527 | 0.0887 | 5.71 |
| 3 | 1.30 | 1.8154 | 1.9937 | 0.1784 | 8.95 |
| 4 | 1.40 | 2.2874 | 2.6117 | 0.3244 | 12.42 |
| 5 | 1.50 | 2.9278 | 3.4904 | 0.5625 | 16.12 |

On a employé cette formule pour calculer la solution exacte $y(x_n)$ dans le tableau. \square

L'exemple suivant illustre les limites de la méthode d'Euler. On verra à la sous-section suivante des méthodes plus précises que celle d'Euler.

EXEMPLE 10.2. Approcher la solution du problème à valeur initiale

$$y'(x) = 2xy, \quad y(1) = 1, \quad (10.6)$$

sur l'intervalle $1 \leq x \leq 1.5$ par la méthode d'Euler avec $h = 0.1$.

RÉSOLUTION. On a, comme à l'exemple précédent,

$$x_0 = 1, \quad x_N = 1.5, \quad y_0 = 1, \quad x_n = x_0 + hn = 1 + 0.1n, \quad N = \frac{1.5 - 1}{0.1} = 5.$$

Avec $f(x, y) = 2xy$, la méthode d'Euler est

$$y_{n+1} = y_n + 0.1 \times 2(1 + 0.1n)y_n, \quad y_0 = 1,$$

pour $n = 0, 1, 2, 3, 4$. Les résultats numériques sont dans le tableau 10.2. L'erreur relative montre que notre approximations n'est pas très bonne. \square

DÉFINITION 10.2. Soit la récurrence

$$y_{n+1} = y_n + h\phi(x_n, y_n). \quad (10.7)$$

On définit l'**erreur de méthode** par l'expression

$$\tau_{n+1} = \frac{1}{h} [y(x_{n+1}) - y(x_n)] - \phi(x_n, y(x_n)) \quad \text{pour } n = 0, 1, 2, \dots, N-1.$$

La méthode (10.7) est d'**ordre** k si $|\tau_j| \leq Mh^k$ pour une constante M et pour tout j .

On trouve une définition équivalente à la section 10.4.

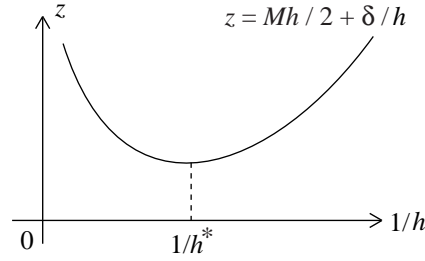


FIGURE 10.1. La courbe des erreurs de méthode et d'arrondi en fonction de $1/h$.

EXEMPLE 10.3. D'après la définition 10.2, l'erreur de la méthode d'Euler est

$$\tau_{n+1} = \frac{1}{h} [y(x_{n+1}) - y(x_n)] - f(x_n, y(x_n)) = \frac{h}{2} y''(\xi_n)$$

pour un certain ξ_n entre x_n et x_{n+1} . Si

$$M = \max_{x_0 \leq x \leq x_N} |y''(x)|,$$

alors $|\tau_n| \leq \frac{h}{2} M$ pour tout n . Donc, la méthode d'Euler est d'ordre un.

REMARQUE 10.1. En général, on ne peut pas affirmer qu'en prenant h suffisamment petit on puisse obtenir le degré de précision désiré, c'est-à-dire que y_n soit aussi près de $y(x_n)$ qu'on le désire. Quand h décroît l'erreur de méthode décroît, mais quand le nombre de pas croît, le nombre d'opérations arithmétiques croît, et par conséquent, les erreurs d'arrondi croissent comme le montre la figure 10.1.

Par exemple, soit y_n la valeur calculée de $y(x_n)$ en (10.4). Posons

$$e_n = y(x_n) - y_n, \quad \text{pour } n = 0, 1, \dots, N.$$

Si

$$|e_0| < \delta_0$$

et δ est la précision des calculs, alors on peut montrer que

$$|e_n| \leq \frac{1}{L} \left(\frac{Mh}{2} + \frac{\delta}{h} \right) \left(e^{L(x_n - x_0)} - 1 \right) + \delta_0 e^{L(x_n - x_0)},$$

où L est la constante de Lipschitz définie dans le théorème 10.1,

$$M = \max_{x_0 \leq x \leq x_N} |y''(x)|,$$

et $h = (x_N - x_0)/N$.

On voit que l'expression

$$z(h) = \frac{Mh}{2} + \frac{\delta}{h}$$

d'abord décroît, puis croît quand $1/h$ croît, comme le montre la figure 10.1. Le 1er terme, $Mh/2$, provient de l'erreur de méthode et le second terme, δ/h , provient des erreurs d'arrondi.

TABLE 10.3. Résultats numériques pour l'exemple 10.4.

| n | x_n | y_n^P | y_n^C | $y(x_n)$ | Erreur absolue | Erreur relative |
|-----|-------|---------|---------|----------|----------------|-----------------|
| 0 | 1.00 | | 1.0000 | 1.0000 | 0.0000 | 0.00 |
| 1 | 1.10 | 1.200 | 1.2320 | 1.2337 | 0.0017 | 0.14 |
| 2 | 1.20 | | 1.5479 | 1.5527 | 0.0048 | 0.31 |
| 3 | 1.30 | | 1.9832 | 1.9937 | 0.0106 | 0.53 |
| 4 | 1.40 | | 2.5908 | 2.6117 | 0.0209 | 0.80 |
| 5 | 1.50 | | 3.4509 | 3.4904 | 0.0344 | 1.13 |

10.2.2. Méthode d'Euler améliorée. La méthode d'Euler améliorée emploie la moyenne des pentes aux deux bouts du pas. On exprime la méthode sous forme de prédicteur et correcteur:

$$y_{n+1}^P = y_n^C + hf(x_n, y_n^C),$$

$$y_{n+1}^C = y_n^C + \frac{1}{2}h [f(x_n, y_n^C) + f(x_{n+1}, y_{n+1}^P)].$$

La méthode est d'ordre 2.

EXEMPLE 10.4. Approcher la solution du problème à valeur initiale de l'exemple 10.2,

$$y'(x) = 2xy, \quad y(1) = 1,$$

sur l'intervalle $1 \leq x \leq 1.5$ par la méthode d'Euler améliorée avec $h = 0.1$.

RÉSOLUTION. On a

$$x_n = x_0 + hn = 1 + 0.1n, \quad n = 0, 1, \dots, 5.$$

Le prédicteur et le correcteur

$$y_0^C = 1,$$

$$y_{n+1}^P = y_n^C + 0.2x_n y_n,$$

$$y_{n+1}^C = y_n^C + 0.1(x_n y_n^C + x_{n+1} y_{n+1}^P)$$

calculent les valeurs prédites y_{n+1}^P et les valeurs corrigées y_{n+1}^C pour $n = 0, 1, \dots, 4$. Les résultats numériques sont dans le tableau 10.3. Ces résultats sont beaucoup plus précis que ceux du tableau 10.2 obtenus par la méthode d'Euler. \square

Dans la suite on va obtenir des méthodes plus précises d'ordre supérieur à 2.

10.3. Méthodes de Runge–Kutta explicites d'ordre 2, 3 et 4

Les méthodes de Runge–Kutta sont à un pas mais elles sont à plusieurs étages.

10.3.1. Méthodes de Runge–Kutta d'ordre 2. Les méthodes de Runge–Kutta explicites à 2 étages sont données par la formule (à gauche) et le tableau de Butcher (à droite):

$$k_1 = hf(x_n, y_n)$$

$$k_2 = hf(x_n + c_2h, y_n + a_{21}k_1)$$

$$y_{n+1} = y_n + b_1k_1 + b_2k_2$$

| | | |
|-----------|----------------------|-------------|
| | c | A |
| k_1 | 0 | 0 |
| k_2 | c_2 | a_{21} 0 |
| y_{n+1} | b^T | b_1 b_2 |

Dans un tableau de Butcher, les composantes du vecteur \mathbf{c} sont les incréments de x_n et les entrées de la matrice A sont les multiplicateurs des pentes approchées, qui, après multiplication par le pas h , deviennent les incréments de y_n . Les composantes du vecteur \mathbf{b} sont les poids qui multiplient les k_j . On a ajouté la colonne tout à fait à gauche du tableau pour aider le lecteur.

On dérive les méthodes de Runge–Kutta du second ordre à deux étages en développant la solution exacte $y(x_{n+1})$, k_1 , k_2 et y_{n+1} autour de x_n par le théorème de Taylor avec reste d'ordre h^3 et l'on pose $h = x_{n+1} - x_n$.

Première étape : Développement de $y(x_{n+1})$ en x_n :

$$\begin{aligned} y(x_{n+1}) &= y(x_n + h) \\ &= y(x_n) + y'(x_n)(x_{n+1} - x_n) + \frac{1}{2}y''(x_n)(x_{n+1} - x_n)^2 \\ &\quad + \frac{1}{6}y'''(\xi_n)(x_{n+1} - x_n)h^3 \\ &= y(x_n) + y'(x_n)h + \frac{1}{2}y''(x_n)h^2 + \frac{1}{6}y'''(\xi_n)h^3 \end{aligned} \quad (10.8)$$

pour un certain ξ_n entre x_n et x_{n+1} et $n = 0, 1, \dots, N-1$. Si l'on dérive l'équation différentielle

$$y'(x) = f(x, y(x)),$$

par rapport à x , on obtient des expressions pour $y'(x_n)$ et $y''(x_n)$:

$$\begin{aligned} y'(x_n) &= f(x_n, y(x_n)), \\ y''(x_n) &= \frac{d}{dx} f(x, y(x)) \Big|_{x=x_n} \\ &= f_x(x_n, y(x_n)) + f_y(x_n, y(x_n)) f(x_n, y(x_n)), \end{aligned}$$

qu'on substitue en (10.8) :

$$\begin{aligned} y(x_{n+1}) &= y(x_n) + f(x_n, y(x_n))h \\ &\quad + \frac{1}{2} \left[f_x(x_n, y(x_n)) + f_y(x_n, y(x_n))f(x_n, y(x_n)) \right] h^2 \\ &\quad + \frac{1}{6} y'''(\xi_n)h^3. \end{aligned} \quad (10.9)$$

Deuxième étape : Développement de tous les termes de la formule de Runge–Kutta en $(x_n, y(x_n))$:

$$\begin{aligned} y_{n+1} &= y_n + b_1 f(x_n, y(x_n))h + b_2 f(x_n, y(x_n))h \\ &\quad + b_2 c_2 f_x(x_n, y(x_n))h^2 \\ &\quad + b_2 a_{21} f_y(x_n, y(x_n))f(x_n, y(x_n))h^2 + O(h^3). \end{aligned} \quad (10.10)$$

Troisième étape : Pour que les expressions (10.9) et (10.10) soient égales à l'ordre h^3 , il faut que

$$b_1 + b_2 = 1, \quad b_2 c_2 = 1/2, \quad b_2 a_{21} = 1/2.$$

On a donc 3 équations à 4 inconnues, ce qui donne lieu à une famille de solutions à 1 paramètre.

Voici quelques méthodes de Runge–Kutta du 2^{ème} ordre à 2 étages.

On écrit la méthode d'Euler améliorée comme une méthode de Runge-Kutta du 2ième ordre à 2 étages (gauche) et son tableau de Butcher (droite) appelée méthode de Heun du second ordre :

$$\begin{array}{l}
 k_1 = hf(x_n, y_n) \\
 k_2 = hf(x_n + h, y_n + k_1) \\
 y_{n+1} = y_n + \frac{1}{2}(k_1 + k_2)
 \end{array}
 \qquad
 \begin{array}{c|cc}
 & \mathbf{c} & A \\
 \hline
 k_1 & 0 & 0 \\
 k_2 & 1 & 1 \quad 0 \\
 \hline
 y_{n+1} & \mathbf{b}^T & 1/2 \quad 1/2
 \end{array}$$

La méthode du point-milieu sous forme de Runge-Kutta :

$$\begin{array}{l}
 k_1 = hf(x_n, y_n) \\
 k_2 = hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1\right) \\
 y_{n+1} = y_n + k_2
 \end{array}
 \qquad
 \begin{array}{c|cc}
 & \mathbf{c} & A \\
 \hline
 k_1 & 0 & 0 \\
 k_2 & 1/2 & 1/2 \quad 0 \\
 \hline
 y_{n+1} & \mathbf{b}^T & 0 \quad 1
 \end{array}$$

Une méthode de Heun :

$$\begin{array}{l}
 k_1 = hf(x_n, y_n) \\
 k_2 = hf\left(x_n + \frac{2}{3}h, y_n + \frac{2}{3}k_1\right) \\
 y_{n+1} = y_n + \frac{1}{4}k_1 + \frac{3}{4}k_2
 \end{array}
 \qquad
 \begin{array}{c|cc}
 & \mathbf{c} & A \\
 \hline
 k_1 & 0 & 0 \\
 k_2 & 2/3 & 2/3 \quad 0 \\
 \hline
 y_{n+1} & \mathbf{b}^T & 1/4 \quad 3/4
 \end{array}$$

10.3.2. Méthodes de Runge-Kutta d'ordre 3. Voici deux méthodes bien connues de Heun et de Kutta du 3ième ordre à 3 étages et leurs tableaux de Butcher.

$$\begin{array}{c|ccc}
 & \mathbf{c} & & A \\
 \hline
 k_1 & 0 & 0 & \\
 k_2 & 1/3 & 1/3 & 0 \\
 k_3 & 2/3 & 0 & 2/3 \quad 0 \\
 \hline
 y_{n+1} & \mathbf{b}^T & 1/4 & 0 \quad 3/4
 \end{array}$$

Tableau de Butcher de la méthode de Heun du 3ième ordre.

$$\begin{array}{c|ccc}
 & \mathbf{c} & & A \\
 \hline
 k_1 & 0 & 0 & \\
 k_2 & 1/2 & 1/2 & 0 \\
 k_3 & 1 & -1 & 2 \quad 0 \\
 \hline
 y_{n+1} & \mathbf{b}^T & 1/6 & 2/3 \quad 1/6
 \end{array}$$

Tableau de Butcher de la méthode de Kutta du 3ième ordre.

10.3.3. Méthodes de Runge-Kutta d'ordre 4. La méthode de Runge-Kutta (classique) d'ordre 4, notée RK4, est une méthode explicite très populaire.

Par le théorème de Taylor avec $h = x_{n+1} - x_n$:

$$\begin{aligned}
 y(x_{n+1}) = y(x_n) + y'(x_n)h + \frac{y''(x_n)}{2!}h^2 + \frac{y^{(3)}(x_n)}{3!}h^3 + \frac{y^{(4)}(x_n)}{4!}h^4 \\
 + \frac{y^{(5)}(\xi_n)}{5!}h^5
 \end{aligned}$$

pour ξ_n entre x_n et x_{n+1} et $n = 0, 1, \dots, N - 1$. On obtient un méthode d'ordre 4 en suivant les étapes utilisées pour la méthode du second ordre. On détermine les poids : b_1, b_2, b_3, b_4 , les incréments en x : c_2, c_3, c_4 , et les incréments en y : a_{21}, a_{32}, a_{43} , pour que

$$y'(x_n)h + \frac{y''(x_n)}{2!}h^2 + \frac{y^{(3)}(x_n)}{3!}h^3 + \frac{y^{(4)}(x_n)}{4!}h^4 + O(h^5)$$

soit égal à

$$b_1k_1 + b_2k_2 + b_3k_3 + b_4k_4 + O(h^5),$$

où

$$\begin{aligned} k_1 &= hf(x_n, y_n), \\ k_2 &= hf(x_n + c_2h, y_n + a_{21}k_1), \\ k_3 &= hf(x_n + c_3h, y_n + a_{32}k_2), \\ k_4 &= hf(x_n + c_4h, y_n + a_{43}k_3). \end{aligned}$$

Ceci suit des relations

$$\begin{aligned} y'(x_n) &= f(x_n, y(x_n)), \\ y''(x_n) &= \frac{d}{dx} f(x, y(x))|_{t=x_n} \\ &= f_x(x_n, y(x_n)) + f_y(x_n, y(x_n)) f(x_n, y(x_n)), \dots, \end{aligned}$$

et une application du théorème de Taylor pour les fonctions de 2 variables. On omet les longs calculs.

On obtient la méthode classique de Runge–Kutta d'ordre 4 et son tableau de Butcher :

$$\begin{aligned} k_1 &= hf(x_n, y_n) \\ k_2 &= hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1\right) \\ k_3 &= hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2\right) \\ k_4 &= hf(x_n + h, y_n + k_3) \\ y_{n+1} &= y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned}$$

| | | | | | |
|-----------|----------------------|----------|-----|-----|-----|
| | c | A | | | |
| k_1 | 0 | 0 | | | |
| k_2 | 1/2 | 1/2 | 0 | | |
| k_3 | 1/2 | 0 | 1/2 | 0 | |
| k_4 | 1 | 0 | 0 | 1 | 0 |
| y_{n+1} | b^T | 1/6 | 2/6 | 2/6 | 1/6 |

L'exemple suivant montre que RK4 résout mieux (10.6) que les méthodes précédentes.

EXEMPLE 10.5. Approcher la solution du problème à valeur initiale de l'exemple 10.2,

$$y'(x) = 2xy, \quad y(1) = 1,$$

sur l'intervalle $1 \leq x \leq 1.5$ par la méthode RK4 avec $h = 0.1$.

RÉSOLUTION. On a $f(x, y) = 2xy$ et

$$x_n = 1.0 + 0.1n, \quad \text{pour } n = 0, 1, \dots, 5.$$

On démarre avec $y_0 = 1.0$. Alors

$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

TABLE 10.4. Résultats numériques pour l'exemple 10.5.

| x_n | y_n | $y(x_n)$ | Erreur absolue | Erreur relative |
|-------|--------|----------|----------------|-----------------|
| 1.00 | 1.0000 | 1.0000 | 0.0000 | 0.0 |
| 1.10 | 1.2337 | 1.2337 | 0.0000 | 0.0 |
| 1.20 | 1.5527 | 1.5527 | 0.0000 | 0.0 |
| 1.30 | 1.9937 | 1.9937 | 0.0000 | 0.0 |
| 1.40 | 2.6116 | 2.6117 | 0.0001 | 0.0 |
| 1.50 | 3.4902 | 3.4904 | 0.0002 | 0.0 |

où

$$k_1 = 0.1 \times 2(1.0 + 0.1n)y_n,$$

$$k_2 = 0.1 \times 2(1.05 + 0.1n)(y_n + k_1/2),$$

$$k_3 = 0.1 \times 2(1.05 + 0.1n)(y_n + k_2/2),$$

$$k_4 = 0.1 \times 2(1.0 + 0.1(n+1))(y_n + k_3),$$

et $n = 0, 1, 2, 3, 4$. Les résultats numériques sont dans le tableau 10.4. \square

EXEMPLE 10.6. Soit le problème à valeur initiale :

$$y' = (y - x - 1)^2 + 2, \quad y(0) = 1.$$

Calculer y_4 par RK4 avec $h = 0.1$.

RÉSOLUTION. La solution est dans le tableau.

| n | x_n | y_n | Valeur exacte $y(x_n)$ | Erreur globale $y(x_n) - y_n$ |
|-----|-------|---------------|---------------------------|----------------------------------|
| 0 | 0.0 | 1.000 000 000 | 1.000 000 000 | 0.000 000 000 |
| 1 | 0.1 | 1.200 334 589 | 1.200 334 672 | 0.000 000 083 |
| 2 | 0.2 | 1.402 709 878 | 1.402 710 036 | 0.000 000 157 |
| 3 | 0.3 | 1.609 336 039 | 1.609 336 250 | 0.000 000 181 |
| 4 | 0.4 | 1.822 792 993 | 1.822 793 219 | 0.000 000 226 |

\square

EXEMPLE 10.7. Approcher à 6 décimales près la solution du problème à valeur initiale

$$y' = x + \arctan y, \quad y(0) = 0,$$

sur l'intervalle $0 \leq x \leq 1$ par la méthode RK4 avec $h = 0.01$. Imprimer la solution à toutes les 10 valeurs et tracer la solution numérique..

RÉSOLUTION. **La solution par Matlab numérique.**— Le fichier M exp10_7 pour l'exemple 10.7 est

```
function yprime = exp10_7(x,y); % Example 12.7.
yprime = x+atan(y);
```

Le programme applique la méthode RK4 :

```

clear
h = 0.01; x0= 0; xf= 1; y0 = 0;
n = ceil((xf-x0)/h); % nombre de pas
%
count = 2; print_time = 10; % quand ecrire les donnees de sortie
x = x0; y = y0; % initialiser x et y
output = [0 x0 y0];
for i=1:n
    k1 = h*exp10_7(x,y);
    k2 = h*exp10_7(x+h/2,y+k1/2);
    k3 = h*exp10_7(x+h/2,y+k2/2);
    k4 = h*exp10_7(x+h,y+k3);
    z = y + (1/6)*(k1+2*k2+2*k3+k4);
    x = x + h;
    if count > print_time
        output = [output; i x z];
        count = count - print_time;
    end
    y = z;
count = count + 1;
end
output
save output % pour tracer le graphique

```

La commande `output` imprime n , x , et y .

| n | x | y |
|----------|--------|--------|
| 0 | 0 | 0 |
| 10.0000 | 0.1000 | 0.0052 |
| 20.0000 | 0.2000 | 0.0214 |
| 30.0000 | 0.3000 | 0.0499 |
| 40.0000 | 0.4000 | 0.0918 |
| 50.0000 | 0.5000 | 0.1486 |
| 60.0000 | 0.6000 | 0.2218 |
| 70.0000 | 0.7000 | 0.3128 |
| 80.0000 | 0.8000 | 0.4228 |
| 90.0000 | 0.9000 | 0.5531 |
| 100.0000 | 1.0000 | 0.7040 |

La commande suivante imprime les données de sortie.

```

load output;
subplot(2,2,1); plot(output(:,2),output(:,3));
title('Plot of solution y_n for Example 12.7');
xlabel('x_n'); ylabel('y_n');

```

□

À l'exemple suivant on résout le système de deux équations de van der Pol par RK4 et par le code `ode23` de Matlab (V. paragraphe 10.7.1). On compare les 2 solutions sur graphique.

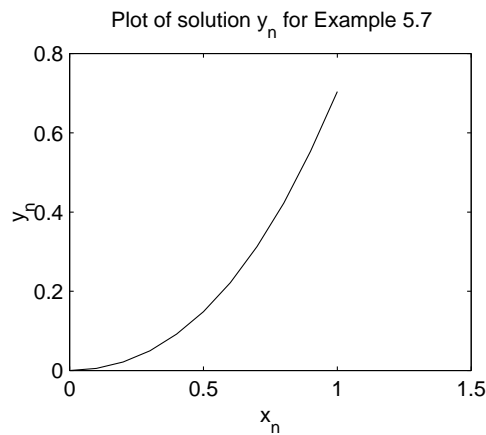


FIGURE 10.2. Tracé de la solution de l'exemple 10.7.

EXEMPLE 10.8. Résoudre l'équation du second ordre de van der Pol

$$y'' + (y^2 - 1)y' + y = 0, \quad y(0) = 0, \quad y'(0) = 0.25, \quad (10.11)$$

sur $0 \leq x \leq 20$ par RK4. Imprimer la solution aux 10 valeurs et tracer la solution. De plus, résoudre (10.11) par le code `ode23` et tracer la solution.

RÉSOLUTION. On réécrit le problème (10.11) sous la forme d'un système de 2 équations du 1er ordre en posant $y_1 = y$ et $y_2 = y'_1$,

$$\begin{aligned} y'_1 &= y_2, \\ y'_2 &= y_2(1 - y_1^2) - y_1, \end{aligned}$$

avec les conditions initiales $y_1(0) = 0$ et $y_2(0) = 0.25$.

Le programme en MATLAB appelle la fonction M `exp1vdp.m` :

```
function yprime = exp1vdp(t,y); % Exemple 10.8.
yprime = [y(2); y(2).*(1-y(1).^2)-y(1)]; % systeme de van der Pol
```

Le programme applique RK4 sur l'équation différentielle définie par le fichier M `exp1vdp.m`:

```
clear
h = 0.1; t0= 0; tf= 21; % pas d'integration, temps initial et final
y0 = [0 0.25]'; % conditions initiales
n = ceil((xf-t0)/h); % nombre de pas

count = 2; print_control = 10; % quand imprimer la donnee
t = t0; y = y0; % initialiser t et y
output = [t0 y0']; % 1ere ligne de la matrice de valeurs a imprimer
w = [t0, y0']; % 1ere ligne de la matrice de valeurs a tracer
for i=1:n
    k1 = h*exp1vdp(x,y);          k2 = h*exp1vdp(x+h/2,y+k1/2);
    k3 = h*exp1vdp(x+h/2,y+k2/2); k4 = h*exp1vdp(x+h,y+k3);
    z = y + (1/6)*(k1+2*k2+2*k3+k4);
    t = t + h;
```

```

    if count > print_control
        output = [output; t z']; % incrementer la matrice des valeurs a imprimer
        count = count - print_control;
    end
    y = z;
    w = [w; t z']; % incrementer la matrice des valeurs a tracer
    count = count + 1;
end
[output(1:11,:) output(12:22,:)] % imprimer la solution numerique
save w % garder en memoire la matrice pour tracer la solution

```

La commande output imprime les valeurs de t , y_1 et y_2 .

| t | y(1) | y(2) | t | y(1) | y(2) |
|---------|---------|---------|---------|---------|---------|
| 0 | 0 | 0.2500 | 11.0000 | -1.9923 | -0.2797 |
| 1.0000 | 0.3586 | 0.4297 | 12.0000 | -1.6042 | 0.7195 |
| 2.0000 | 0.6876 | 0.1163 | 13.0000 | -0.5411 | 1.6023 |
| 3.0000 | 0.4313 | -0.6844 | 14.0000 | 1.6998 | 1.6113 |
| 4.0000 | -0.7899 | -1.6222 | 15.0000 | 1.8173 | -0.5621 |
| 5.0000 | -1.6075 | 0.1456 | 16.0000 | 0.9940 | -1.1654 |
| 6.0000 | -0.9759 | 1.0662 | 17.0000 | -0.9519 | -2.6628 |
| 7.0000 | 0.8487 | 2.5830 | 18.0000 | -1.9688 | 0.3238 |
| 8.0000 | 1.9531 | -0.2733 | 19.0000 | -1.3332 | 0.9004 |
| 9.0000 | 1.3357 | -0.8931 | 20.0000 | 0.1068 | 2.2766 |
| 10.0000 | -0.0939 | -2.2615 | 21.0000 | 1.9949 | 0.2625 |

Les commandes suivantes tracent le graphique de la solution.

```

load w % charger les valeurs pour produire le graphique
subplot(2,2,1); plot(w(:,1),w(:,2)); % tracer la solution par RK4
title('Solution y_n par RK4 pour l'exemple 12.8');
xlabel('t_n'); ylabel('y_n');

```

Maintenant on emploie le code ode23.

```

load w % charger les valeurs pour produire le graphique
v = [0 21 -3 3]; % definir les axes t et y
subplot(2,2,1);
plot(w(:,1),w(:,2)); % tracer la solution obtenue par RK4
axis(v);
title('Solution y_n par RK4 pour l'exemple 12.8');
xlabel('t_n'); ylabel('y_n');
subplot(2,2,2);
[t,y] = ode23('exp1vdp',[0 21], y0);
plot(x,y(:,1)); % tracer la solution obtenue par ode23
axis(v);
title('Solution y_n par ode23 pour l'exemple 12.8');
xlabel('t_n'); ylabel('y_n');

```

Le code ode23 produit les vecteurs : t de 144 nœuds inégalement espacés et les solutions correspondantes $y(1)$ et $y(2)$. La partie gauche de la figure 10.3 est la

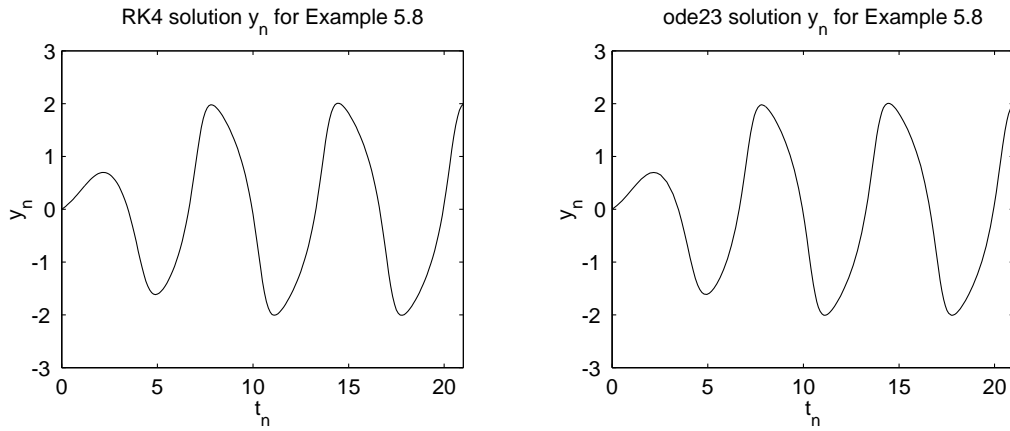


FIGURE 10.3. Graphique de la solution numérique de l'exemple 10.8.

solution par RK4 et la partie droite est celle de `ode23`. On voit que les tracés sont identiques. \square

10.4. Convergence des méthodes numériques

Dans cette section et les deux prochaines, on introduit les concepts de convergence, de consistance et de stabilité des méthodes numériques pour les équations différentielles.

On peut écrire les méthodes numériques considérées dans ce chapitre sous la forme générale

$$\sum_{n=0}^k \alpha_j y_{n+j} = h \varphi_f(y_{n+k}, y_{n+k-1}, \dots, y_n, x_n; h), \quad (10.12)$$

où l'indice inférieur f sur φ indique la dépendance de φ sur la fonction $f(x, y)$ de (10.1). On requiert la condition :

$$\varphi_{f \equiv 0}(y_{n+k}, y_{n+k-1}, \dots, y_n, x_n; h) \equiv 0.$$

On remarque que la continuité lipschitzienne de φ par rapport à y_{n+j} , $n = 0, 1, \dots, k$, suit de la continuité lipschitzienne (10.2) sur f .

DÉFINITION 10.3. On dit que la méthode (10.12) est **convergente** si l'on a

$$y_n - y(x_n) \rightarrow 0 \quad \text{quand } h \downarrow 0,$$

pour tout problème (10.1) et pour $nh = x \in [a, b]$.

L'**erreur de méthode** de (10.12) est le résiduel

$$R_{n+k} := \sum_{n=0}^k \alpha_j y(x_{n+j}) - h \varphi_f(y(x_{n+k}), y(x_{n+k-1}), \dots, y(x_n), x_n; h). \quad (10.13)$$

DÉFINITION 10.4. On dit que la méthode (10.12) est **consistante** si l'on a

$$\frac{1}{h} R_{n+k} \rightarrow 0 \quad \text{quand } h \downarrow 0,$$

pour tout problème (10.1) et pour $nh = x \in [a, b]$.

DÉFINITION 10.5. La méthode (10.12) est **zéro-stable** si les zéros du polynôme caractéristique

$$\sum_{n=0}^k \alpha_j r^{n+j}$$

sont dans le disque unité $|r| \leq 1$ et les zéros sur le cercle unité $|r| = 1$ sont simples.

On a finalement le théorème fondamental suivant.

THÉORÈME 10.2. Une méthode est **convergente** quand $h \downarrow 0$ si et seulement si elle est zéro-stable et consistante.

Toutes les méthodes numériques considérées dans ce chapitre sont convergentes.

10.5. Stabilité absolue des méthodes numériques

Dans les applications on emploie des méthodes numériques consistantes et zéro-stables avec un pas petit mais strictement positif.

Pour $n = 0, 1, 2, \dots$, soit y_{n+1} la solution numérique de (10.1) en $x = x_{n+1}$, et $y^{[n]}(x_{n+1})$ la solution exacte du problème **local**:

$$y' = f(x, y), \quad y(x_n) = y_n. \quad (10.14)$$

On note

$$\varepsilon_{n+1} = y_{n+1} - y^{[n]}(x_{n+1}) \quad (10.15)$$

l'**erreur locale** d'une méthode numérique.

Si $y(x) \in C^{p+1}[x_0, x_N]$ et

$$\varepsilon_{n+1} \approx C_{p+1} h_{n+1}^{p+1} y^{(p+1)}(x_n) + O(h_{n+1}^{p+2}), \quad (10.16)$$

on dit que l'erreur locale est d'ordre $p+1$ et C_{p+1} est la constante de l'erreur de la méthode. L'erreur globale d'une méthode consistante et zéro-stable est d'ordre p si l'erreur locale est d'ordre $p+1$. Dans ce cas, on dit que la méthode est d'ordre p . D'après la définition 10.4, une méthode d'ordre $p \geq 1$ est consistante.

On applique la méthode (10.12), avec $h > 0$ petit, sur l'équation linéaire de test

$$y' = \lambda y, \quad \Re \lambda < 0. \quad (10.17)$$

La **région de stabilité absolue**, R , est la région du plan complexe \widehat{h} , où $\widehat{h} = h\lambda$, pour laquelle la solution numérique y_n de (10.17) tend vers zéro, quand n tend vers l'infini.

La région de stabilité absolue de la méthode d'Euler explicite est le disque de rayon 1 et de centre $(-1, 0)$ (V. la courbe $k = 1$ dans la figure 10.7). La région de stabilité absolue de la méthode d'Euler rétrograde est l'extérieur du disque de rayon 1 et de centre $(1, 0)$; donc elle comprend tout le demi-plan gauche (V. la courbe $k = 1$ dans la figure 10.10).

La région de stabilité absolue, R , d'une méthode explicite est très approximativement un disque ou une cardioïde dans le demi-plan gauche (la cardioïde s'étend sur le plan droit avec un point de rebroussement à l'origine). Le bord de R coupe l'axe des réels en α , tel que $-\infty < \alpha < 0$, et en l'origine. On appelle $[\alpha, 0]$ l'**intervalle de stabilité absolue**. Pour les méthodes à coefficients réels R est symétrique par rapport à l'axe des réels. Toutes les méthodes considérées dans ce

travail ont des coefficients réels; donc il suffit de montrer la partie supérieure de R dans les figures 10.7, 10.8 et 10.10.

La région de stabilité R des méthode implicites s'étend jusqu'à $\alpha = -\infty$. L'angle soustendu à l'origine par R dans le demi-plan gauche est ordinairement plus petit pour les méthodes d'ordre supérieur (V. la figure 10.10).

Si la région R n'inclut pas tout l'axe des réels négatifs, c'est-à-dire $-\infty < \alpha < 0$, alors l'inclusion

$$h\lambda \in R$$

restreint le pas :

$$\alpha \leq h \operatorname{Re} \lambda \implies 0 < h \leq \frac{\alpha}{\operatorname{Re} \lambda}.$$

En pratique, on veut employer un pas h suffisamment petit pour assurer la précision de la solution numérique d'après (10.15)–(10.16), mais pas trop petit.

10.6. Stabilité des méthodes de Runge–Kutta

Il existe de méthodes de Runge–Kutta explicites d'ordre $p = s$ à s étages pour $s = 1, 2, 3, 4$. Pour les méthodes d'ordre 5 il faut au moins 6 étages.

Si l'on applique une méthode de Runge–Kutta method sur l'équation de test

$$y' = \lambda y, \quad \Re \lambda < 0,$$

avec solution $y(x) \rightarrow 0$ quand $t \rightarrow \infty$, on obtient une équation aux différences à un pas :

$$y_{n+1} = Q(\hat{h})y_n, \quad \hat{h} = h\lambda,$$

où $Q(\hat{h})$ est la **fonction de stabilité** de la méthode. On voit que $y_n \rightarrow 0$ quand $n \rightarrow \infty$ si et seulement si

$$|Q(\hat{h})| < 1. \quad (10.18)$$

La méthode est **absolument stable** pour les valeurs de \hat{h} dans le plan complexe qui satisfont (10.18); ces valeurs forment la **région de stabilité absolue** de la méthode. On peut montrer que la fonction de stabilité des méthodes de Runge–Kutta explicites d'ordre $p = s$ à s étage pour $s = 1, 2, 3, 4$ est

$$R(\hat{h}) = \frac{y_{n+1}}{y_n} = 1 + \hat{h} + \frac{1}{2!} \hat{h}^2 + \cdots + \frac{1}{s!} \hat{h}^s.$$

La région de stabilité absolue R d'une méthode de Runge–Kutta d'ordre $k = s$ à s étages pour $s = 1, 2, 3, 4$ forme l'intérieur de la k ème région dont la partie supérieure est dans la figure 10.4. Le point le plus à gauche de R est $\alpha = -2, -2, -2.51, -2.78$ respectivement pour les méthodes d'ordre $s = 1, 2, 3, 4$.

La fonction M de Matlab, téléchargée du site <ftp://ftp.cs.cornell.edu/pub/cv> implémente les méthodes de Runge–Kutta à pas fixe d'ordre 1 à 5.

```
function [tvals,yvals] = FixedRK(fname,t0,y0,h,k,n)
%
% Produces approximate solution to the initial value problem
%
%      y'(t) = f(t,y(t))      y(t0) = y0
%
% using a strategy that is based upon a k-th order
% Runge-Kutta method. Stepsize is fixed.
%
% Pre:  fname = string that names the function f.
```

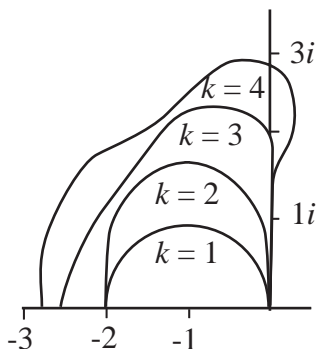


FIGURE 10.4. Région de stabilité absolue des méthodes Runge-Kutta explicites d'ordre $k = s$ à s étages.

```

%      t0 = initial time.
%      y0 = initial condition vector.
%      h = stepsize.
%      k = order of method. (1<=k<=5).
%      n = number of steps to be taken,
%
% Post: tvals(j) = t0 + (j-1)h, j=1:n+1
%       yvals(:j) = approximate solution at t = tvals(j), j=1:n+1
%
tc = t0;
yc = y0;
tvals = tc;
yvals = yc;
fc = feval(fname,tc,yc);
for j=1:n
    [tc,yc,fc] = RKstep(fname,tc,yc,fc,h,k);
    yvals = [yvals yc ];
    tvals = [tvals tc];
end
function [tnew,ynew,fnew] = RKstep(fname,tc,yc,fc,h,k)
%
% Pre:  fname is a string that names a function of the form f(t,y)
%       where t is a scalar et y is a column d-vecteur.
%
%       yc is an approximate solution to y'(t) = f(t,y(t)) at t=tc.
%
%       fc = f(tc,yc).
%
%       h is the time step.
%
%       k is the order of the Runge-Kutta method used, 1<=k<=5.
%
```

```

% Post: tnew=tc+h, ynew is an approximate solution at t=tnew, et
%       fnew = f(tnew,ynew).

if k==1
    k1 = h*fc;
    ynew = yc + k1;

elseif k==2
    k1 = h*fc;
    k2 = h*feval(fname,tc+h,yc+k1);
    ynew = yc + (k1 + k2)/2;

elseif k==3
    k1 = h*fc;
    k2 = h*feval(fname,tc+(h/2),yc+(k1/2));
    k3 = h*feval(fname,tc+h,yc-k1+2*k2);
    ynew = yc + (k1 + 4*k2 + k3)/6;

elseif k==4
    k1 = h*fc;
    k2 = h*feval(fname,tc+(h/2),yc+(k1/2));
    k3 = h*feval(fname,tc+(h/2),yc+(k2/2));
    k4 = h*feval(fname,tc+h,yc+k3);
    ynew = yc + (k1 + 2*k2 + 2*k3 + k4)/6;

elseif k==5
    k1 = h*fc;
    k2 = h*feval(fname,tc+(h/4),yc+(k1/4));
    k3 = h*feval(fname,tc+(3*h/8),yc+(3/32)*k1
                +(9/32)*k2);
    k4 = h*feval(fname,tc+(12/13)*h,yc+(1932/2197)*k1
                -(7200/2197)*k2+(7296/2197)*k3);
    k5 = h*feval(fname,tc+h,yc+(439/216)*k1
                - 8*k2 + (3680/513)*k3 -(845/4104)*k4);
    k6 = h*feval(fname,tc+(1/2)*h,yc-(8/27)*k1
                + 2*k2 -(3544/2565)*k3 + (1859/4104)*k4 - (11/40)*k5);
    ynew = yc + (16/135)*k1 + (6656/12825)*k3 +
                (28561/56430)*k4 - (9/50)*k5 + (2/55)*k6;

end
tnew = tc+h;
fnew = feval(fname,tnew,ynew);

```

10.7. Paires de méthodes de Runge-Kutta

Jusqu'à présent on a seulement considéré un pas h constant. Dans les applications il est avantageux de faire varier h : on augmente h quand $y(x)$ varie peu et on le réduit quand $y(x)$ varie rapidement. On considère maintenant les méthodes à pas variable.

Les paires de méthodes Runge–Kutta imbriquées d'ordre p et $p+1$ contrôlent l'erreur locale et le pas au moyen de la différence $y_{n+1} - \widehat{y}_{n+1}$ entre les solutions d'ordre supérieur et d'ordre inférieur. Certaines paires sont aussi dotées d'un interpolant pour calculer la solution numérique entre les nœuds et, dans certains cas, contrôler le pas.

10.7.1. La paire ode23 de RK à 4 étages de Matlab. Le code `ode23` est une paire de méthodes Runge–Kutta imbriquées d'ordre 2 et 3 à 4 étages avec contrôle de l'erreur. La méthode avance la solution de y_n à y_{n+1} au moyen de la méthode d'ordre 3 (appelé extrapolation locale) et contrôle le pas au moyen de la solution du second ordre. La formule suivante présente les 4 étages :

$$\begin{aligned} k_1 &= h f(x_n, y_n), \\ k_2 &= h f(x_n + (1/2)h, y_n + (1/2)k_1), \\ k_3 &= h f(x_n + (3/4)h, y_n + (3/4)k_2), \\ k_4 &= h f(x_n + h, y_n + (2/9)k_1 + (1/3)k_2 + (4/9)k_3), \end{aligned}$$

Les 3 premiers étages servent au calcul de la solution en x_{n+1} :

$$y_{n+1} = y_n + \frac{2}{9}k_1 + \frac{1}{3}k_2 + \frac{4}{9}k_3,$$

et les 4 étages servent à estimer l'erreur locale :

$$E = -\frac{5}{72}k_1 + \frac{1}{12}k_2 + \frac{1}{9}k_3 - \frac{1}{8}k_4.$$

On remarque que cette méthode est à 3 étages parce que le 1er étage en x_{n+1} est le 4ème en x_n , c'est-à-dire $k_1^{[n+1]} = k_4^{[n]}$. On dit que ces méthodes sont FSAL (First Same As Last).

L'interpolant naturel de la paire `ode23` est un polynôme d'Hermite de degré 3 à 2 points qui interpole y_n et $f(x_n, y_n)$ en $x = x_n$ et y_{n+1} et $f(x_{n+1}, y_{n+1})$ en $x = x_{n+1}$.

Voici le tableau de Butcher de la paire `ode23` :

| | c | A | | | |
|---------------------|---|-----------------|----------------|---------------|----------------|
| k_1 | 0 | 0 | | | |
| k_2 | $\frac{1}{2}$ | $\frac{1}{2}$ | | | |
| k_3 | $\frac{3}{4}$ | 0 | $\frac{3}{4}$ | | |
| k_4 | 1 | $\frac{2}{9}$ | $\frac{1}{3}$ | $\frac{4}{9}$ | |
| y_{n+1} | \mathbf{b}^T | $\frac{2}{9}$ | $\frac{1}{3}$ | $\frac{4}{9}$ | 0 |
| \widehat{y}_{n+1} | $\widehat{\mathbf{b}}^T$ | $\frac{7}{24}$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{8}$ |
| EE | $\mathbf{b}^T - \widehat{\mathbf{b}}^T$ | $-\frac{5}{72}$ | $\frac{1}{12}$ | $\frac{1}{9}$ | $-\frac{1}{8}$ |

EXEMPLE 10.9. Soit le problème à valeur initiale :

$$y' = xy + 1, \quad y(0) = 1.$$

Calculer $y(0.1)$ et $y(0.2)$ à 5 décimales près par `ode23` de Matlab et estimer l'erreur locale en ces points.

RÉSOLUTION. Le second membre de l'équation différentielle est

$$f(x, y) = xy + 1.$$

Pour $n = 0$:

$$\begin{aligned}k_1 &= 0.1 \times 1 = 0.1 \\k_2 &= 0.1 \times (0.05 \times 1.05 + 1) = 0.105\,25 \\k_3 &= 0.1 \times (0.75 \times 1.078\,937\,5 + 1) = 0.108\,092\,031\,25 \\k_4 &= 0.1 \times (0.1 \times 1.105\,346\,458\,333\,33 + 1) = 0.111\,053\,464\,583\,33 \\y_1 &= 1.105\,346\,458\,333\,33\end{aligned}$$

et

$$\text{Estimation de l'erreur locale} = -4.506\,848\,958\,333\,448\text{e} - 05$$

Pour $n = 1$:

$$\begin{aligned}k_1 &= 0.111\,053\,464\,583\,33 \\k_2 &= 0.117\,413\,097\,859\,37 \\k_3 &= 0.120\,884\,609\,930\,24 \\k_4 &= 0.124\,457\,783\,972\,15 \\y_2 &= 1.222\,889\,198\,607\,30\end{aligned}$$

et

$$\text{Estimation de l'erreur locale} = -5.322\,100\,094\,209\,102\text{e} - 05$$

Pour résoudre le problème à valeur initiale et tracer la solution sur $[0, 1]$, par la commande `ode23` de Matlab, on écrit le fichier fonction M `exp10_9.m` :

```
function yprime = exp10_9(x,y)
yprime = x.*y+1;
```

et l'on emploie les commandes

```
clear
xspan = [0 1]; y0 = 1; % xspan et valeur initiale
[x,y] = ode23(@exp10_9,xspan,y0);
subplot(2,2,1); plot(x,y); xlabel('x'); ylabel('y');
title('Solution de l\'equation de l\'exemple 12.9');
print -deps2 Figexp10_9 % imprimer la figure dans le fichier Fig.exp10.9
```

□

Le solveur `ode23` de Matlab implémente la paire de Runge-Kutta (2,3) explicite de Bogacki et Shampine appelée BS23. Elle emploie un interpolant “gratuit” d’ordre 3. Elle fait l’extrapolation locale, c’est-à-dire elle avance la solution avec la méthode d’ordre supérieur, c’est-à-dire du 3ème ordre.

10.7.2. La paire de Dormand-Prince DP(5,4)7M à 7 étages avec interpolant. On présente la paire de Dormand-Prince DP(5,4)7M à sept étages avec estimation de l’erreur locale et interpolant dans son tableau de Butcher. Le nombre 5 dans DP(5,4)7M indique qu’on avance la solution avec y_{n+1} d’ordre 5 (procédure appelée *extrapolation locale*). Le nombre 4 indique que la solution \hat{y}_{n+1} d’ordre 4 sert à obtenir une estimation de l’erreur locale au moyen de la différence $y_{n+1} - \hat{y}_{n+1}$. De fait, on ne calcule pas \hat{y}_{n+1} ; on emploie plutôt les coefficients de

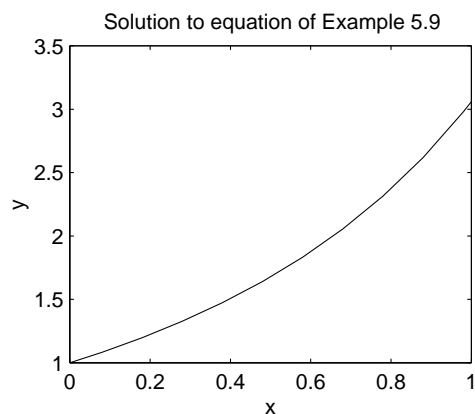


FIGURE 10.5. Tracé de la solution de l'exemple 10.9.

la ligne $b^T - \widehat{b}^T$ pour obtenir l'estimation de l'erreur locale. Le nombre 7 indique que la méthode est à 7 étages. La lettre M indique que la constante C_6 dans le premier terme de l'erreur a été minimisé tout en maintenant la stabilité de la méthode. Une méthode d'ordre 5 requiert au moins 6 étages. Le 7ème étage est requis pour l'interpolant qui utilise la dernière ligne du tableau.

La méthode à 7 étages se réduit à 6 étages si le pas est accepté puisque $k_1^{[n+1]} = k_7^{[n]}$; de fait le vecteur ligne b^T est identique à la 7ème ligne qui correspond à k_7 . C'est donc une méthode FSAL ("First Same As Last") puisque la 1ère ligne et la dernière sont identiques.

L'intervalle de stabilité absolue de la paire DP(5,4)7M est approximativement $(-3.3, 0)$ (V. la figure. 10.6).

La matrice A dans le tableau de Butcher d'une méthode Rung-Kutta explicite est triangulaire inférieure stricte. Les méthodes semi-explicites ont une matrice triangulaire inférieure. Sinon, la méthode est implicite. Résoudre une méthode semi-explicite pour le vecteur solution y_{n+1} du système est moins coûteux que de résoudre une méthode implicite.

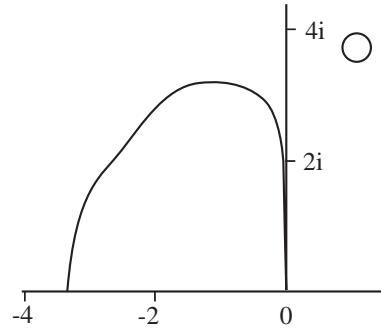


FIGURE 10.6. Région de stabilité absolue de la paire DP(5,4)7M.

| c | | A | | | | | | |
|-----------------|-------------------------------------|----------------------------|-----------------------|--------------------------|--------------------------|------------------------------|-----------------------|---------------------|
| k_1 | 0 | 0 | | | | | | |
| k_2 | $\frac{1}{5}$ | $\frac{1}{5}$ | 0 | | | | | |
| k_3 | $\frac{3}{10}$ | $\frac{3}{40}$ | $\frac{9}{40}$ | 0 | | | | |
| k_4 | $\frac{4}{5}$ | $\frac{44}{45}$ | $-\frac{56}{15}$ | $\frac{32}{9}$ | 0 | | | |
| k_5 | $\frac{8}{9}$ | $\frac{19372}{6561}$ | $-\frac{25360}{2187}$ | $\frac{64448}{6561}$ | $-\frac{212}{729}$ | 0 | | |
| k_6 | 1 | $\frac{9017}{3168}$ | $-\frac{355}{33}$ | $\frac{46732}{5247}$ | $\frac{49}{176}$ | $-\frac{5103}{18656}$ | 0 | |
| k_7 | 1 | $\frac{35}{384}$ | 0 | $\frac{500}{1113}$ | $\frac{125}{192}$ | $-\frac{2187}{6784}$ | $\frac{11}{84}$ | |
| \hat{y}_{n+1} | $\hat{\mathbf{b}}^T$ | $\frac{5179}{57600}$ | 0 | $\frac{7571}{16695}$ | $\frac{393}{640}$ | $-\frac{92097}{339200}$ | $\frac{187}{2100}$ | $\frac{1}{40}$ |
| y_{n+1} | \mathbf{b}^T | $\frac{35}{384}$ | 0 | $\frac{500}{1113}$ | $\frac{125}{192}$ | $-\frac{2187}{6784}$ | $\frac{11}{84}$ | 0 |
| EEL | $\mathbf{b}^T - \hat{\mathbf{b}}^T$ | $\frac{71}{57600}$ | 0 | $-\frac{71}{16695}$ | $\frac{71}{1920}$ | $-\frac{17253}{339200}$ | $\frac{22}{525}$ | $-\frac{1}{40}$ |
| $y_{n+0.5}$ | | $\frac{5783653}{57600000}$ | 0 | $\frac{466123}{1192500}$ | $-\frac{41347}{1920000}$ | $\frac{16122321}{339200000}$ | $-\frac{7117}{20000}$ | $\frac{183}{10000}$ |

(10.19)

La paire DP(5,4)7M d'ordre 5 et 4 à 7 étages avec interpolant.

Les méthodes de Runge-Kutta se fondent sur une idée ingénieuse. La solution unique d'un problème initial bien posé est une courbe bien définie dans \mathbb{R}^{n+1} . Mais les erreurs de méthode et d'arrondi perturbent cette courbe qui devient affectée par les courbes voisines. Donc les méthodes de Runge-Kutta essaient d'obtenir de l'information sur la famille de courbes voisines et non pas uniquement sur la solution unique. Ceci se voit clairement dans le cas des méthodes de Runge-Kutta explicites.

Le solveur `ode45` de Matlab est une implémentation de la paire Runge-Kutta (5,4) explicite de Dormand et Prince appelée RK5(4)7FM, DOPRI5, DP(4,5) et DP54. Elle emploie un interpolant "gratuit" d'ordre 4 fourni privément par Dormand et Prince. Le solveur emploie l'extrapolation locale.

On trouve dans l'article *The MATLAB ODE Suite* de L. F. Shampine et M. W. Reichelt, SIAM Journal on Scientific Computing, **18**(1), 1997, beaucoup de détails sur les solveurs Matlab `ode23`, `ode45` etc.

10.7.3. La paire RKF(4,5) à 6 étages de Fehlberg. La paire Runge–Kutta–Fehlberg RKF(4,5) à 6 étages avec estimation de l'erreur locale avance le pas de y_n à y_{n+1} par une méthode d'ordre 4 et contrôle l'erreur par la différence $y_{n+1} - \hat{y}_{n+1}$ où la valeur auxiliaire \hat{y}_{n+1} est obtenue par une méthode d'ordre 5. On peut lire la méthode dans son tableau de Butcher. On obtient l'estimation de l'erreur locale au moyen de la dernière ligne.

$$\begin{array}{r|ccccccc}
 k_1 & 0 & & & & & & \\
 k_2 & \frac{1}{4} & & & & & & \\
 k_3 & \frac{3}{8} & & & & & & \\
 k_4 & \frac{12}{13} & & & & & & \\
 k_5 & 1 & & & & & & \\
 k_6 & \frac{1}{2} & & & & & & \\
 \hline
 y_{n+1} & \mathbf{b}^T & & & & & & \\
 \hat{y}_{n+1} & \hat{\mathbf{b}}^T & & & & & & \\
 \hat{\mathbf{b}}^T - \mathbf{b}^T & & & & & & &
 \end{array}
 \quad (10.20)$$

La paire Runge–Kutta–Fehlberg RKF(4,5) d'ordre 4 et 5 à 6 étages.

L'intervalle de stabilité absolue de la paire RKF(4,5) est approximativement $(-3.78, 0)$. RKF45 minimise la constante C_5 de la méthode d'ordre 4 et c'est elle qui fait avancer la solution de y_n à y_{n+1} , c'est-à-dire, sans extrapolation locale. L'algorithme suivant implémente la méthode.

ALGORITHME 10.1. Soit y_0 la condition initiale et $h > 0$. Supposons que l'approximation y_n de $y(x_n)$ satisfasse l'inégalité $|y(x_n) - y_n| < \epsilon$ où ϵ est la précision recherchée.

- (1) Calculer 2 approximations pour y_{n+1} : la 1ère par la méthode d'ordre 4 :

$$y_{n+1} = y_n + \left(\frac{25}{216}k_1 + \frac{1408}{2565}k_3 + \frac{2197}{4104}k_4 - \frac{1}{5}k_5 \right), \quad (10.21)$$

et la 2ème par la méthode d'ordre 5 :

$$\hat{y}_{j+1} = y_n + \left(\frac{16}{135}k_1 + \frac{6656}{12825}k_3 + \frac{28561}{56430}k_4 - \frac{9}{50}k_5 + \frac{2}{55}k_6 \right), \quad (10.22)$$

où

$$k_1 = hf(x_n, y_n),$$

$$k_2 = hf(x_n + h/4, y_n + k_1/4),$$

$$k_3 = hf(x_n + 3h/8, y_n + 3k_1/32 + 9k_2/32),$$

$$k_4 = hf(x_n + 12h/13, y_n + 1932k_1/2197 - 7200k_2/2197 + 7296k_3/2197),$$

$$k_5 = hf(x_n + h, y_n + 439k_1/216 - 8k_2 + 3680k_3/513 + 845k_4/4104),$$

$$k_6 = hf(x_n + h/2, y_n - 8k_1/27 + 2k_2 + 3544k_3/2565 + 1859k_4/4104 - 11k_5/40).$$

- (2) Si $|\hat{y}_{j+1} - y_{n+1}| < \epsilon h$, on accepte l'approximation y_{n+1} de $y(x_{n+1})$. On remplace alors h par qh où

$$q = [\epsilon h / (2|\hat{y}_{j+1} - y_{n+1}|)]^{1/4}$$

et l'on revient à l'étape (1) pour calculer l'approximation y_{j+2} .

(3) Si $|\hat{y}_{j+1} - y_{n+1}| \geq \epsilon h$, on remplace h par qh où

$$q = [\epsilon h / (2|\hat{y}_{j+1} - y_{n+1}|)]^{1/4}$$

et l'on revient à l'étape (1) pour calculer l'approximation suivante de y_{n+1} .

On peut montrer que l'erreur locale de (10.21) est approximativement

$$|\hat{y}_{j+1} - y_{n+1}|/h.$$

À l'étape (2), on veut que l'erreur soit inférieure à ϵh en module pour avoir $|y(x_n) - y_n| < \epsilon$ pour tout j (et en particulier $|y(x_N) - y_f| < \epsilon$). La formule pour calculer q en (2) et (3) (et par conséquent une nouvelle valeur de h) est dérivée de la relation entre les erreurs locales de (10.21) et de (10.22).

RKF(4,5) surestime l'erreur de la solution d'ordre 4 parce que la constante de l'erreur est minimisée. La méthode RKV de Verner qui suit corrige cette faille.

10.7.4. La paire RKV(5,6) à 8 étages de Verner. La paire Runge-Kutta-Verner RKV(5,6) d'ordre 5 et 6 à 8 étages se lit dans son tableau de Butcher. On remarque que l'ordre 6 requiert au moins 8 étages. La méthode essaie d'avoir une erreur globale proportionnelle à la tolérance visée par l'utilisateur. Cette méthode est idéale pour les problèmes non raides pour lesquels le calcul de la fonction $f(x, y)$ est peu coûteux et si l'on ne cherche pas la solution en un grand nombre de points très rapprochés (par exemple pour fin de graphique).

| | c | A | | | | | | |
|-----------------|----------------------|-----------------------|-------------------|------------------------|---------------------|-----------------------|----------------|--------------------------------------|
| k_1 | 0 | 0 | | | | | | |
| k_2 | $\frac{1}{6}$ | $\frac{1}{6}$ | 0 | | | | | |
| k_3 | $\frac{4}{15}$ | $\frac{4}{75}$ | $\frac{16}{75}$ | 0 | | | | |
| k_4 | $\frac{2}{3}$ | $\frac{5}{6}$ | $-\frac{8}{3}$ | $\frac{5}{2}$ | 0 | | | |
| k_5 | $\frac{5}{6}$ | $-\frac{165}{64}$ | $\frac{55}{6}$ | $-\frac{425}{64}$ | $\frac{85}{96}$ | 0 | | |
| k_6 | 1 | $\frac{12}{5}$ | -8 | $\frac{4015}{612}$ | $-\frac{11}{36}$ | $\frac{88}{255}$ | 0 | (10.23) |
| k_7 | $\frac{1}{15}$ | $-\frac{8263}{15000}$ | $\frac{124}{75}$ | $-\frac{643}{680}$ | $-\frac{81}{250}$ | $\frac{2484}{10625}$ | 0 | |
| k_8 | 1 | $\frac{3501}{1720}$ | $-\frac{300}{43}$ | $\frac{297275}{52632}$ | $-\frac{319}{2322}$ | $\frac{24068}{84065}$ | 0 | $\frac{3850}{26703}$ |
| y_{n+1} | \mathbf{b}^T | $\frac{13}{160}$ | 0 | $\frac{2375}{5984}$ | $\frac{5}{16}$ | $\frac{12}{85}$ | $\frac{3}{44}$ | |
| \hat{y}_{n+1} | $\hat{\mathbf{b}}^T$ | $\frac{3}{40}$ | 0 | $\frac{875}{2244}$ | $\frac{23}{72}$ | $\frac{264}{1955}$ | 0 | $\frac{125}{11592}$ $\frac{43}{616}$ |

Paire de Runge-Kutta-Verner RKV(5,6) d'ordre 5 et 6.

10.8. Méthodes prédicteurs-correcteurs multipas

10.8.1. Méthodes multipas générales. Soit le problème à valeur initiale

$$y' = f(x, y), \quad y(a) = \eta. \quad (10.24)$$

On suppose que la fonction $f(x, y)$ est continue en x et lipschitzienne en y sur la bande $[a, b] \times (-\infty, \infty)$. Alors, par le théorème 10.1, la solution exacte $y(x)$, existe et est unique sur $[a, b]$.

On cherche une solution numérique approchée $\{y_n\}$ aux nœuds $x_n = a + nh$, $n = 1, 2, \dots, N$ avec le pas h et $N = (b - a)/h$. On emploie une *méthode multipas* à k pas :

$$\sum_{j=0}^k \alpha_j y_{n+j} = h \sum_{j=0}^k \beta_j f_{n+j}, \quad (10.25)$$

où $y_n \approx y(x_n)$ et $f_n := f(x_n, y_n)$. On normalise la méthode par la condition $\alpha_k = 1$ et l'on exige que le nombre de pas soit exactement k en supposant que

$$(\alpha_0, \beta_0) \neq (0, 0).$$

On calcule k valeurs de départ y_0, y_1, \dots, y_{k-1} , disons, par une méthode de Runge-Kutta du même ordre.

La méthode est *explicite* si $\beta_k = 0$. Dans ce cas on obtient y_{n+1} directement. La méthode est *implicite* si $\beta_k \neq 0$. Dans ce cas il faut calculer y_{n+k} par la récurrence :

$$y_{n+k}^{[s+1]} = h\beta_k f(x_{n+k}, y_{n+k}^{[s]}) + g, \quad y_{n+k}^{[0]} \text{ arbitraire}, \quad s = 0, 1, \dots, \quad (10.26)$$

où la fonction

$$g = g(x_n, \dots, x_{n+k-1}, y_0, \dots, y_{n+k-1})$$

ne contient que des valeurs connues. La récurrence (10.26) converge quand $s \rightarrow \infty$, si $0 \leq M < 1$ où M est la constante de Lipschitz du second membre de (10.26) par rapport à y_{n+k} . Si L est la constante de Lipschitz de $f(x, y)$ en y , alors

$$M := Lh|\beta_k| < 1 \quad (10.27)$$

et la convergence suit de l'inégalité

$$h < \frac{1}{L|\beta_k|}.$$

Si l'on applique (10.25) sur l'équation de test

$$y' = \lambda y, \quad \Re \lambda < 0,$$

avec solution $y(x) \rightarrow 0$ quand $t \rightarrow \infty$, on trouve que la solution numérique $y_n \rightarrow 0$ quand $n \rightarrow \infty$ si, avec $\widehat{h} = \lambda h$, les zéros $r_s(\widehat{h})$ du polynôme de stabilité

$$\pi(r, \widehat{h}) := \sum_{n=0}^k (\alpha_j - \widehat{h}\beta_j)r^j$$

satisfont $|r_s(\widehat{h})| \leq 1$, $s = 1, 2, \dots, k$ et $|r_s(\widehat{h})| < 1$ si le zéro est multiple. Dans ce cas on dit que la méthode linéaire multipas (10.25) est **absolument stable** pour \widehat{h} . La **région de stabilité absolue** R dans le plan complexe est l'ensemble des valeurs \widehat{h} pour lesquelles la méthode est absolument stable.

10.8.2. Méthodes linéaires multipas d'Adams-Bashforth-Moulton.

Les méthodes linéaires à k pas d'Adams-Bashforth (AB) (explicites) et d'Adams-Moulton (AM) (implicites) sont, respectivement, de la forme

$$y_{n+1} - y_n = h \sum_{j=0}^{k-1} \beta_j^* f_{n+j-k+1}, \quad y_{n+1} - y_n = h \sum_{j=0}^k \beta_j f_{n+j-k+1}$$

Les tableaux 10.5 et 10.6 présentent les coefficients des méthodes AB et AM de 1 à 6 pas. Dans le tableau, les coefficients sont à diviser par d , k dénote le nombre de pas, p désigne l'ordre et C_{p+1}^* et C_{p+1} sont les constantes de l'erreur des méthodes.

TABLE 10.5. Coefficients des méthodes d'Adams–Bashforth de 1 à 6 pas.

| β_5^* | β_4^* | β_3^* | β_2^* | β_1^* | β_0^* | d | k | p | C_{p+1}^* |
|-------------|-------------|-------------|-------------|-------------|-------------|------|-----|-----|-------------|
| | | | | | 1 | 1 | 1 | 1 | 1/2 |
| | | | | 3 | -1 | 2 | 2 | 2 | 5/12 |
| | | | 23 | -16 | 5 | 12 | 3 | 3 | 3/8 |
| | | 55 | -59 | 37 | -9 | 24 | 4 | 4 | 251/720 |
| | 1901 | -2774 | 1616 | -1274 | 251 | 720 | 5 | 5 | 95/288 |
| 4277 | -7923 | 9982 | -7298 | 2877 | -475 | 1440 | 6 | 6 | 19087/60480 |

TABLE 10.6. Coefficients des méthodes d'Adams–Moulton de 1 à 6 pas.

| β_5 | β_4 | β_3 | β_2 | β_1 | β_0 | d | k | p | C_{p+1} |
|-----------|-----------|-----------|-----------|-----------|-----------|------|-----|-----|------------|
| | | | | 1 | 1 | 2 | 1 | 2 | -1/12 |
| | | | 5 | 8 | -1 | 12 | 2 | 3 | -1/24 |
| | | 9 | 19 | -5 | 1 | 24 | 3 | 4 | -19/720 |
| | 251 | 646 | -264 | 106 | -19 | 720 | 4 | 5 | -3/160 |
| 475 | 1427 | -798 | 482 | -173 | 27 | 1440 | 5 | 6 | -863/60480 |

La région de stabilité absolue des méthodes d'Adams–Bashforth et d'Adams–Moulton à k pas d'ordre $k = 1, 2, 3, 4$ est l'intérieur des régions dont la partie supérieure se trouve dans la figure 10.7. La région de stabilité absolue de la méthode d'Adams–Bashforth d'ordre 3 s'étend sur une petite région triangulaire dans le demi-plan droit. La région de stabilité absolue de la méthode d'Adams–Moulton d'ordre 1 recouvre tout le demi-plan gauche.

En pratique, on emploie une méthode AB comme **prédicteur** pour obtenir la valeur prédite y_{n+1}^* en x_{n+1} . Cette valeur est insérée dans le second membre d'une méthode AM employée comme **correcteur** pour obtenir la valeur corrigée y_{n+1} . Une telle combinaison s'appelle un prédicteur-correcteur ABM. Si le prédicteur et le correcteur sont du même ordre on peut employer l'estimation de Milne du terme principal de l'erreur de méthode :

$$\epsilon_{n+1} \approx \frac{C_{p+1}}{C_{p+1}^* - C_{p+1}} (y_{n+1} - y_{n+1}^*).$$

Ce procédé, appelé *extrapolation locale*, améliore la solution y_{n+1} d'ordre supérieur par l'addition de l'estimateur de l'erreur locale :

$$y_{n+1} + \frac{C_{p+1}}{C_{p+1}^* - C_{p+1}} (y_{n+1} - y_{n+1}^*).$$

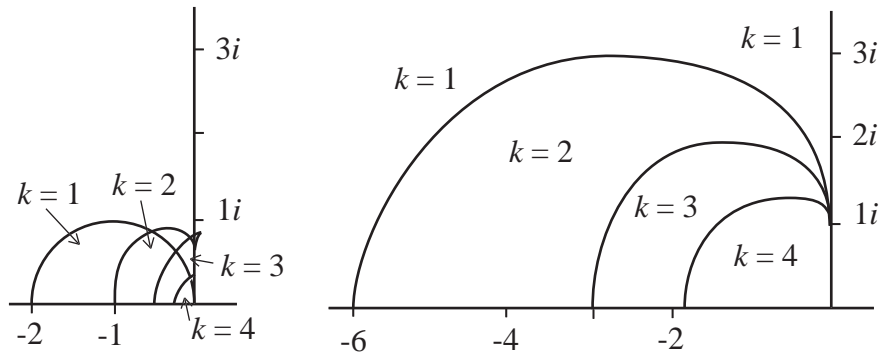


FIGURE 10.7. Régions de stabilité absolue des méthodes d'ordre k d'Adams–Bashforth (à gauche) et d'Adams–Moulton (à droite).

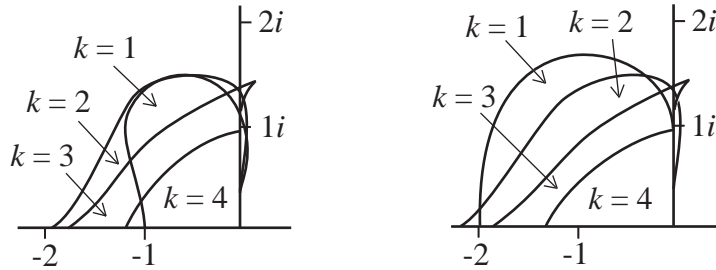


FIGURE 10.8. Régions de stabilité absolue des méthodes d'Adams–Bashforth–Moulton d'ordre k en mode PECE (à gauche) et en mode PECLE (à droite).

Les régions de stabilité absolue des paires Adams–Bashforth–Moulton d'ordre k pour $k = 1, 2, 3, 4$ en mode *prédicteur-évaluation-corrector-évaluation*, noté PECE, sont symétriques par rapport à l'axe des réels. Leur partie supérieure se trouve à l'intérieur des régions dans la partie gauche de la figure 10.8. De même, les régions de stabilité absolue de ces paires en mode PECLE, où L veut dire extrapolation locale, se trouvent dans la partie droite de la figure 10.8.

10.8.3. Méthodes d'Adams–Bashforth–Moulton d'ordre 3 et 4. On considère, comme 1er exemple, la méthode d'Adams–Bashforth–Moulton d'ordre 3 :

$$y_{n+1}^P = y_n^C + \frac{h}{12} (23f_n^C - 16f_{n-1}^C + 5f_{n-2}^C), \quad f_k^C = f(x_k, y_k^C), \quad (10.28)$$

$$y_{n+1}^C = y_n^C + \frac{h}{12} (5f_{n+1}^P + 8f_n^C - f_{n-1}^C), \quad f_k^P = f(x_k, y_k^P), \quad (10.29)$$

avec estimation de l'erreur locale

$$\text{Err.} \approx -\frac{1}{10} [y_{n+1}^C - y_{n+1}^P]. \quad (10.30)$$

EXEMPLE 10.10. Résoudre à 6 décimales près, le problème à valeur initiale

$$y' = x + \sin y, \quad y(0) = 0,$$

par la méthode d'Adams–Bashforth–Moulton d'ordre 3 sur l'intervalle $[0, 2]$ avec $h = 0.2$. On a déjà obtenu les valeurs initiales par une méthode de précision élevée. Estimer l'erreur locale à chaque pas au moyen de la formule (10.30).

RÉSOLUTION. La solution est dans le tableau suivant.

| n | x_n | Départ y_n^C | Prédite y_n^P | Corrigée y_n^C | $10^5 \times$ Erreur locale en y_n^C $\approx -(y_n^C - y_n^P) \times 10^4$ |
|-----|-------|-------------------|--------------------|---------------------|--|
| 0 | 0.0 | 0.000 000 0 | | | |
| 1 | 0.2 | 0.021 404 7 | | | |
| 2 | 0.4 | 0.091 819 5 | | | |
| 3 | 0.6 | | 0.221 260 | 0.221 977 | – 7 |
| 4 | 0.8 | | 0.423 703 | 0.424 064 | – 4 |
| 5 | 1.0 | | 0.710 725 | 0.709 623 | 11 |
| 6 | 1.2 | | 1.088 004 | 1.083 447 | 46 |
| 7 | 1.4 | | 1.542 694 | 1.533 698 | 90 |
| 8 | 1.6 | | 2.035 443 | 2.026 712 | 87 |
| 9 | 1.8 | | 2.518 039 | 2.518 431 | – 4 |
| 10 | 2.0 | | 2.965 994 | 2.975 839 | –98 |

□

On considère, comme 2ème exemple, la méthode d'Adams–Bashforth–Moulton d'ordre 4 très connue :

Le *prédicteur d'Adams–Bashforth* et le *correcteur d'Adams–Moulton* d'ordre 4 sont

$$y_{n+1}^P = y_n^C + \frac{h}{24} (55f_n^C - 59f_{n-1}^C + 37f_{n-2}^C - 9f_{n-3}^C) \quad (10.31)$$

et

$$y_{n+1}^C = y_n^C + \frac{h}{24} (9f_{n+1}^P + 19f_n^C - 5f_{n-1}^C + f_{n-2}^C), \quad (10.32)$$

où

$$f_n^C = f(x_n, y_n^C) \quad \text{et} \quad f_n^P = f(x_n, y_n^P).$$

On obtient les valeurs de départ par une méthode de Runge–Kutta ou autrement.

On contrôle l'erreur locale par l'estimation

$$C_5 h^5 y^{(5)}(x_{n+1}) \approx -\frac{19}{270} [y_{n+1}^C - y_{n+1}^P]. \quad (10.33)$$

On garde en mémoire un certain nombre de valeurs antérieures de y_n et de f_n pour augmenter le pas si l'erreur locale est petite devant la tolérance choisie. Si l'erreur locale est trop forte devant la tolérance, on coupe le pas en deux au moyen des deux formules suivantes :

$$y_{n-1/2} = \frac{1}{128} (35y_n + 140y_{n-1} - 70y_{n-2} + 28y_{n-3} - y_{n-4}), \quad (10.34)$$

$$y_{n-3/2} = \frac{1}{162} (-y_n + 24y_{n-1} + 54y_{n-2} - 16y_{n-3} + 3y_{n-4}). \quad (10.35)$$

L'intervalle de stabilité absolue de la paire d'Adams–Bashforth–Moulton d'ordre 4 en mode PECE est $(-1.25, 0)$, c'est-à-dire, la méthode n'amplifie pas les erreurs antérieures si le pas h est suffisamment petit pour que

$$-1.25 < h \frac{\partial f}{\partial y} < 0, \quad \text{où} \quad \frac{\partial f}{\partial y} < 0.$$

EXEMPLE 10.11. Soit le problème à valeur initiale

$$y' = x + y, \quad y(0) = 0.$$

Calculer la solution en $x = 2$ par la méthode d'Adams–Bashforth–Moulton d'ordre 4 avec $h = 0.2$. Employer la méthode de Runge–Kutta d'ordre 4 pour calculer les valeurs de départ. Faire les calculs à 5 décimales près et calculer l'erreur globale au moyen de la solution exacte $y(x) = e^x - x - 1$.

RÉSOLUTION. Les résultats sont dans le tableau suivant.

| n | x_n | Départ y_n^C | Prédites y_n^P | Corrigées y_n^C | Exactes $y(x_n)$ | Erreurs: $10^6 \times$ $(y(x_n) - y_n^C)$ |
|-----|-------|-------------------|---------------------|----------------------|---------------------|--|
| 0 | 0.0 | 0.000 000 | | | 0.000 000 | 0 |
| 1 | 0.2 | 0.021 400 | | | 0.021 403 | 3 |
| 2 | 0.4 | 0.091 818 | | | 0.091 825 | 7 |
| 3 | 0.6 | 0.222 107 | | | 0.222 119 | 12 |
| 4 | 0.8 | | 0.425 361 | 0.425 529 | 0.425 541 | 12 |
| 5 | 1.0 | | 0.718 066 | 0.718 270 | 0.718 282 | 12 |
| 6 | 1.2 | | 1.119 855 | 1.120 106 | 1.120 117 | 11 |
| 7 | 1.4 | | 1.654 885 | 1.655 191 | 1.655 200 | 9 |
| 8 | 1.6 | | 2.352 653 | 2.353 026 | 2.353 032 | 6 |
| 9 | 1.8 | | 3.249 190 | 3.249 646 | 3.249 647 | 1 |
| 10 | 2.0 | | 4.388 505 | 4.389 062 | 4.389 056 | -6 |

On voit que la méthode est stable puisque l'erreur ne croît pas. □

EXEMPLE 10.12. Résoudre le problème à valeur initiale

$$y' = \arctan x + \arctan y, \quad y(0) = 0,$$

à 6 décimales près par la méthode d'Adams–Bashforth–Moulton d'ordre 3 sur l'intervalle $[0, 2]$ avec $h = 0.2$. Employer RK4 pour calculer les valeurs de départ. Estimer l'erreur locale à chaque pas par la formule (10.30).

RÉSOLUTION. **La solution par Matlab numérique.**— Le fichier `M exp10_12` pour l'exemple 10.12 est

```
function yprime = exp10_12(x,y); % Exemple 10.12.
yprime = atan(x)+atan(y);
```

On calcule les 4 valeurs de départ par RK4 et la condition initiale.

```
clear
h = 0.2; x0= 0; xf= 2; y0 = 0;
n = ceil((xf-x0)/h); % nombre de pas
%
count = 2; print_time = 1; % quand ecrire les donnees de sortie
x = x0; y = y0; % initialiser x et y
```

```

output = [0 x0 y0 0];
%RK4
for i=1:3
    k1 = h*exp10_12(x,y);
    k2 = h*exp10_12(x+h/2,y+k1/2);
    k3 = h*exp10_12(x+h/2,y+k2/2);
    k4 = h*exp10_12(x+h,y+k3);
    z = y + (1/6)*(k1+2*k2+2*k3+k4);
    x = x + h;
    if count > print_time
        output = [output; i x z 0];
        count = count - print_time;
    end
    y = z;
count = count + 1;
end
% ABM4
for i=4:n
    zp = y + (h/24)*(55*exp10_12(output(i,2),output(i,3))-...
        59*exp10_12(output(i-1,2),output(i-1,3))+...
        37*exp10_12(output(i-2,2),output(i-2,3))-...
        9*exp10_12(output(i-3,2),output(i-3,3)) );
    z = y + (h/24)*( 9*exp10_12(x+h,zp)+...
        19*exp10_12(output(i,2),output(i,3))-...
        5*exp10_12(output(i-1,2),output(i-1,3))+...
        exp10_12(output(i-2,2),output(i-2,3)) );
    x = x + h;
    if count > print_time
        errest = -(19/270)*(z-zp);
        output = [output; i x z errest];
        count = count - print_time;
    end
    y = z;
count = count + 1;
end
output
save output % pour imprimer le graphique

```

La commande `output` imprime les valeurs de n , x et y .

| n | x | y | Estimation de l'erreur |
|---|-----|------------------|------------------------|
| 0 | 0 | 0 | 0 |
| 1 | 0.2 | 0.02126422549044 | 0 |
| 2 | 0.4 | 0.08962325332457 | 0 |
| 3 | 0.6 | 0.21103407185113 | 0 |
| 4 | 0.8 | 0.39029787517821 | 0.00001007608281 |
| 5 | 1.0 | 0.62988482479868 | 0.00005216829834 |
| 6 | 1.2 | 0.92767891924367 | 0.00004381671342 |

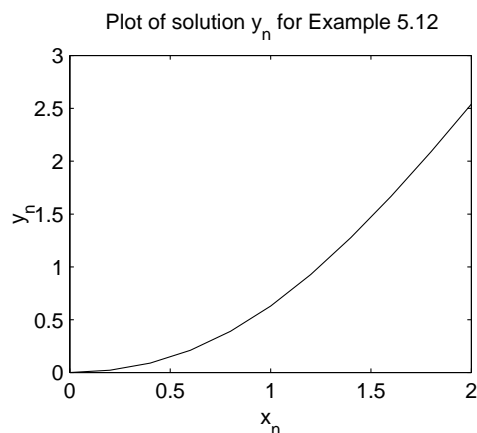


FIGURE 10.9. Tracé de la solution de l'exemple 10.12.

| | | | |
|----|-----|------------------|-------------------|
| 7 | 1.4 | 1.27663327419538 | -0.00003607372725 |
| 8 | 1.6 | 1.66738483675693 | -0.00008228934754 |
| 9 | 1.8 | 2.09110753309673 | -0.00005318684309 |
| 10 | 2.0 | 2.54068815072267 | -0.00001234568256 |

Les commandes suivantes impriment les données de sortie.

```
load output;
subplot(2,2,1); plot(output(:,2),output(:,3));
title('Tracé de la solution y_n pour l'exemple 12.12');
xlabel('x_n'); ylabel('y_n');
```

□

La fonction du fichier M <ftp://ftp.cs.cornell.edu/pub/cv> de Matlab implémente les méthodes d'Adams–Bashforth–Moulton d'ordre 1 à 5 à pas fixe,

```
function [tvals,yvals] = FixedPC(fname,t0,y0,h,k,n)
%
% Produces an approximate solution to the initial value problem
%
%      y'(t) = f(t,y(t))      y(t0) = y0
%
% using a strategy that is based upon a k-th order
% Adams PC method. Stepsize is fixed.
%
% Pre:  fname = string that names the function f.
%       t0 = initial time.
%       y0 = initial condition vecteur.
%       h = stepsize.
%       k = order of method. (1<=k<=5).
%       n = number of steps to be taken,
%
% Post: tvals(j) = t0 + (j-1)h, j=1:n+1
```

```

%      yvals(:j) = approximate solution at t = tvals(j), j=1:n+1
%
      [tvals,yvals,fvals] = StartAB(fname,t0,y0,h,k);
      tc = tvals(k);
      yc = yvals(:,k);
      fc = fvals(:,k);

      for j=k:n
          % Take a step et then update.
          [tc,yPred,fPred,yc,fc] = PCstep(fname,tc,yc,fvals,h,k);
          tvals = [tvals tc];
          yvals = [yvals yc];
          fvals = [fc fvals(:,1:k-1)];
      end
end

```

On obtient les valeurs de départ par le fichier M d'une méthode de Runge–Kutta.

```

function [tvals,yvals,fvals] = StartAB(fname,t0,y0,h,k)
%
% Uses k-th order Runge-Kutta to generate approximate
% solutions to
%          y'(t) = f(t,y(t))   y(t0) = y0
%
% at t = t0, t0+h, ... , t0 + (k-1)h.
%
% Pre:
%     fname is a string that names the function f.
%     t0 is the initial time.
%     y0 is the initial value.
%     h is the step size.
%     k is the order of the RK method used.
%
% Post:
%     tvals = [ t0, t0+h, ... , t0 + (k-1)h].
%     For j =1:k, yvals(:,j) = y(tvals(j)) (approximately).
%     For j =1:k, fvals(:,j) = f(tvals(j),yvals(j)) .
%
      tc = t0;
      yc = y0;
      fc = feval(fname,tc,yc);
      tvals = tc;
      yvals = yc;
      fvals = fc;

      for j=1:k-1
          [tc,yc,fc] = RKstep(fname,tc,yc,fc,h,k);
          tvals = [tvals tc];
          yvals = [yvals yc];
          fvals = [fc fvals];
      end
end

```

end

On trouve le fichier M `Rkstep` à la sous-section 10.6. Le fichier M suivant applique le prédicteur d'Adams–Bashforth.

```
function [tnew,ynew,fnew] = ABstep(fname,tc,yc,fvals,h,k)
%
% Pre: fname is a string that names a function of the form f(t,y)
%       where t is a scalar et y is a column d-vecteur.
%
%       yc is an approximate solution to y'(t) = f(t,y(t)) at t=tc.
%
%       fvals is an d-by-k matrix where fvals(:,i) is an approximation
%       to f(t,y) at t = tc +(1-i)h, i=1:k
%
%       h is the time step.
%
%       k is the order of the AB method used, 1<=k<=5.
%
% Post: tnew=tc+h, ynew is an approximate solution at t=tnew, et
%       fnew = f(tnew,ynew).

if k==1
    ynew = yc + h*fvals;
elseif k==2
    ynew = yc + (h/2)*(fvals*[3;-1]);
elseif k==3
    ynew = yc + (h/12)*(fvals*[23;-16;5]);
elseif k==4
    ynew = yc + (h/24)*(fvals*[55;-59;37;-9]);
elseif k==5
    ynew = yc + (h/720)*(fvals*[1901;-2774;2616;-1274;251]);
end
tnew = tc+h;
fnew = feval(fname,tnew,ynew);
```

Le fichier M suivant applique le correcteur d'Adams–Moulton.

```
function [tnew,ynew,fnew] = AMstep(fname,tc,yc,fvals,h,k)
%
% Pre: fname is a string that names a function of the form f(t,y)
%       where t is a scalar et y is a column d-vecteur.
%
%       yc is an approximate solution to y'(t) = f(t,y(t)) at t=tc.
%
%       fvals is an d-by-k matrix where fvals(:,i) is an approximation
%       to f(t,y) at t = tc +(2-i)h, i=1:k
%
%       h is the time step.
%
%       k is the order of the AM method used, 1<=k<=5.
%
```

```
% Post: tnew=tc+h, ynew is an approximate solution at t=tnew, et
%       fnew = f(tnew,ynew).
```

```
if k==1
    ynew = yc + h*fvals;
elseif k==2
    ynew = yc + (h/2)*(fvals*[1;1]);
elseif k==3
    ynew = yc + (h/12)*(fvals*[5;8;-1]);
elseif k==4
    ynew = yc + (h/24)*(fvals*[9;19;-5;1]);
elseif k==5
    ynew = yc + (h/720)*(fvals*[251;646;-264;106;-19]);
end
tnew = tc+h;
fnew = feval(fname,tnew,ynew);
```

Le fichier M suivant applique le prédicteur-correcteur.

```
function [tnew,yPred,fPred,yCorr,fCorr] = PCstep(fname,tc,yc,fvals,h,k)
%
% Pre:  fname is a string that names a function of the form f(t,y)
%       where t is a scalar et y is a column d-vecteur.
%
%       yc is an approximate solution to y'(t) = f(t,y(t)) at t=tc.
%
%       fvals is an d-by-k matrix where fvals(:,i) is an approximation
%       to f(t,y) at t = tc +(1-i)h, i=1:k
%
%       h is the time step.
%
%       k is the order of the Runge-Kutta method used, 1<=k<=5.
%
% Post: tnew=tc+h,
%       yPred is the predicted solution at t=tnew
%       fPred = f(tnew,yPred)
%       yCorr is the corrected solution at t=tnew
%       fCorr = f(tnew,yCorr).
```

```
[tnew,yPred,fPred] = ABstep(fname,tc,yc,fvals,h,k);
[tnew,yCorr,fCorr] = AMstep(fname,tc,yc,[fPred fvals(:,1:k-1)],h,k);
```

10.8.4. Spécification des méthodes multipas. Le 1er membre d'une méthode d'Adams est de la forme

$$y_{n+1} - y_n.$$

Les méthodes d'Adams–Bashforth sont explicites et celles d'Adams–Moulton sont implicites. On obtient les méthodes d'Adams dans les formules qui suivent en prenant $a = 0$, $b = 0$ et $c = 0$. L'entier k est le nombre de pas de la méthode.

L'entier p est l'ordre de la méthode et la constante C_{p+1} est la constante du 1er terme de l'erreur.

Méthodes explicites

$k = 1 :$

$$\begin{aligned}\alpha_1 &= 1, \\ \alpha_0 &= -1, \quad \beta_0 = 1, \\ p &= 1; \quad C_{p+1} = \frac{1}{2}.\end{aligned}$$

$k = 2 :$

$$\begin{aligned}\alpha_2 &= 1, \\ \alpha_1 &= -1 - a, \quad \beta_1 = \frac{1}{2}(3 - a), \\ \alpha_0 &= a, \quad \beta_0 = \frac{1}{2}(-1 + a), \\ p &= 2; \quad C_{p+1} = \frac{1}{12}(5 + a).\end{aligned}$$

La stabilité absolue limite l'ordre à 2.

$k = 3 :$

$$\begin{aligned}\alpha_3 &= 1, \\ \alpha_2 &= -1 - a, \quad \beta_2 = \frac{1}{12}(23 - 5a - b), \\ \alpha_1 &= a + b, \quad \beta_1 = \frac{1}{3}(-4 - 2a + 2b), \\ \alpha_0 &= -b, \quad \beta_0 = \frac{1}{12}(5 + a + 5b), \\ p &= 3; \quad C_{p+1} = \frac{1}{24}(9 + a + b).\end{aligned}$$

La stabilité absolue limite l'ordre à 3.

$k = 4 :$

$$\begin{aligned}\alpha_4 &= 1, \\ \alpha_3 &= -1 - a, \quad \beta_3 = \frac{1}{24}(55 - 9a - b - c), \\ \alpha_2 &= a + b, \quad \beta_2 = \frac{1}{24}(-59 - 19a + 13b - 19c), \\ \alpha_1 &= -b - c, \quad \beta_1 = \frac{1}{24}(37 + 5a + 13b - 19c), \\ \alpha_0 &= c, \quad \beta_0 = \frac{1}{24}(-9 - a - b - 9c), \\ p &= 4; \quad C_{p+1} = \frac{1}{720}(251 + 19a + 11b + 19c).\end{aligned}$$

La stabilité absolue limite l'ordre à 4.

Méthodes implicites

$k = 1 :$

$$\begin{aligned}\alpha_1 &= 1, \quad \beta_1 = \frac{1}{2}, \\ \alpha_0 &= -1, \quad \beta_0 = \frac{1}{2}, \\ p &= 2; \quad C_{p+1} = -\frac{1}{12}.\end{aligned}$$

$k = 2 :$

$$\begin{aligned} \alpha_2 &= 1, & \beta_2 &= \frac{1}{12}(5 + a), \\ \alpha_1 &= -1 - a, & \beta_1 &= \frac{2}{3}(1 - a), \\ \alpha_0 &= a, & \beta_0 &= \frac{1}{12}(-1 - 5a), \\ \text{Si } a &\neq -1, \quad p = 3; & C_{p+1} &= -\frac{1}{24}(1 + a), \\ \text{Si } a &= -1, \quad p = 4; & C_{p+1} &= -\frac{1}{90}. \end{aligned}$$

$k = 3 :$

$$\begin{aligned} \alpha_3 &= 1, & \beta_3 &= \frac{1}{24}(9 + a + b), \\ \alpha_2 &= -1 - a, & \beta_2 &= \frac{1}{24}(19 - 13a - 5b), \\ \alpha_1 &= a + b, & \beta_1 &= \frac{1}{24}(-5 - 13a + 19b), \\ \alpha_0 &= -b, & \beta_0 &= \frac{1}{24}(1 + a + 9b), \\ p &= 4; & C_{p+1} &= -\frac{1}{720}(19 + 11a + 19b). \end{aligned}$$

La stabilité absolue limite l'ordre à 4.

$k = 4 :$

$$\begin{aligned} \alpha_4 &= 1, & \beta_4 &= \frac{1}{720}(251 + 19a + 11b + 19c), \\ \alpha_3 &= -1 - a, & \beta_3 &= \frac{1}{360}(323 - 173a - 37b - 53c), \\ \alpha_2 &= a + b, & \beta_2 &= \frac{1}{30}(-11 - 19a + 19b + 11c), \\ \alpha_1 &= -b - c, & \beta_1 &= \frac{1}{360}(53 + 37a + 173b - 323c), \\ \alpha_0 &= c, & \beta_0 &= \frac{1}{720}(-19 - 11a - 19b - 251c). \end{aligned}$$

Si $27 + 11a + 11b + 27c \neq 0$, alors

$$p = 5; \quad C_{p+1} = -\frac{1}{1440}(27 + 11a + 11b + 27c).$$

Si $27 + 11a + 11b + 27c = 0$, alors

$$p = 6; \quad C_{p+1} = -\frac{1}{15\,120}(74 + 10a - 10b - 74c).$$

La stabilité absolue limite l'ordre à 6.

Le solveur `ode113` de Matlab est complètement à pas variable. Il implémente la famille d'Adams–Bashforth–Moulton d'ordre 1 à 12 en mode PECE au moyen de différences divisées modifiées. Il utilise des interpolants naturels gratuits et avance la solution par extrapolation locale. On trouve des détails sur cette famille dans l'article *The MATLAB ODE Suite* de L. F. Shampine et M. W. Reichelt, SIAM Journal on Scientific Computing, **18**(1), 1997.

10.9. Systèmes différentiels raides

Dans cette section, on illustre le concept de systèmes d'équations différentielles raides au moyen d'un exemple et l'on mentionne quelques méthodes numériques pour résoudre ces systèmes.

10.9.1. Le phénomène de raideur. Bien que le sens intuitif de **système raide** soit clair chez les spécialistes, il y a beaucoup de controverses sur la définition mathématique de ce concept. L'opinion la plus pragmatique est historiquement la première : une équation est raide si certaines méthodes implicites, comme les méthodes à différentiation rétrograde, sont beaucoup plus performantes que les méthodes explicites pour résoudre cette équation.

Soit le système à n équations différentielles :

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y}),$$

et $\lambda_1, \lambda_2, \dots, \lambda_n$ les valeurs propres de la matrice jacobienne de rang n :

$$J = \frac{\partial \mathbf{f}}{\partial \mathbf{y}} = \left(\frac{\partial f_i}{\partial y_j} \right), \quad i \downarrow 1, \dots, n, \quad j \rightarrow 1, \dots, n, \quad (10.36)$$

où l'on a employé la notation de Nagumo. On suppose que les valeurs propres de J ont la partie réelle négative, $\operatorname{Re} \lambda_j < 0$, et qu'elles sont ordonnées d'après la partie réelle :

$$\operatorname{Re} \lambda_n \leq \dots \leq \operatorname{Re} \lambda_2 \leq \operatorname{Re} \lambda_1 < 0. \quad (10.37)$$

Le rapport de raideur sera utile.

DÉFINITION 10.6. Le **rapport de raideur** du système $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$ est le nombre positif

$$r = \frac{\operatorname{Re} \lambda_n}{\operatorname{Re} \lambda_1}, \quad (10.38)$$

où les valeurs propres de la matrice jacobienne (10.36) du système satisfont les inégalités (10.37).

Le phénomène de raideur a plusieurs visages :

- Un système linéaire à coefficients constants est raide si la partie réelle de toutes les valeurs propre est négative et le rapport de raideur est grand.
- Il y a raideur quand les conditions de stabilité, plutôt que celles de précision, réduisent le pas.
- Il y a raideur quand certaines composantes de la solution décroissent beaucoup plus rapidement que les autres.
- Un système est raide sur un intervalle donné I contenant t si en I les solutions voisines approchent la solution à un taux très grand en comparaison avec le taux de variation de la solution en I .

La définition suivante prend en compte ce qu'on observe en pratique.

DÉFINITION 10.7. Un système est raide si une méthode numérique avec région de stabilité absolue bornée appliquée sur un système différentiel avec conditions initiales quelconques doit employer un pas **excessivement petit** sur un intervalle I même si la solution est lisse.

Les méthodes de Runge–Kutta explicites et les paires prédicteurs-correcteurs, qui de fait sont explicites, ne peuvent résoudre un système raide économiquement encore si elles peuvent le faire. Les méthodes implicites requièrent la solution d'équations nonlinéaires le plus souvent obtenues par une variété de méthodes de Newton. On présente 2 méthodes implicites dans les 2 sous-sections qui suivent.

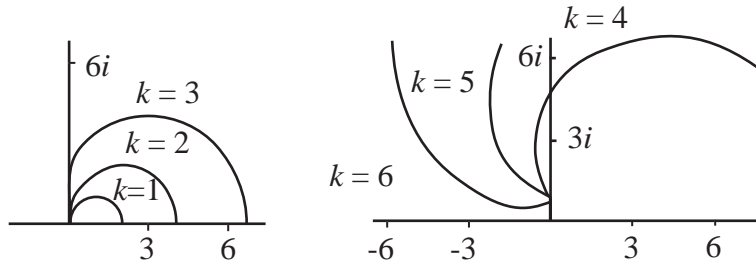


FIGURE 10.10. Régions de stabilité absolue des BDF à $k = 1, 2, \dots, 6$ pas. Ces régions inclut l'axe des réels négatifs.

10.9.2. Formule de différentiation rétrograde. On définit une formule de **différentiation rétrograde** à k pas (“backward differentiation formula”) (BDF) sous forme standard par l’expression

$$\sum_{j=0}^k \alpha_j y_{n+j-k+1} = h \beta_k f_{n+1},$$

où $\alpha_k = 1$. Les BDF sont des méthodes implicites. On trouve dans le tableau 10.7 les BDF de 1 à 6 pas où k est le nombre de pas, p est l’ordre, C_{p+1} est la constante de l’erreur et α est le demi angle soutendu à l’origine par la région de stabilité absolue R .

TABLE 10.7. Coefficients des méthodes BDF.

| k | α_6 | α_5 | α_4 | α_3 | α_2 | α_1 | α_0 | β_k | p | C_{p+1} | α |
|---|------------|--------------------|--------------------|--------------------|--------------------|-------------------|-------------------|------------------|-----|--------------------|----------|
| 1 | | | | | | 1 | -1 | 1 | 1 | 1 | 90° |
| 2 | | | | | 1 | $-\frac{4}{3}$ | $\frac{1}{3}$ | $\frac{2}{3}$ | 2 | $-\frac{2}{9}$ | 90° |
| 3 | | | | 1 | $-\frac{18}{11}$ | $\frac{9}{11}$ | $-\frac{2}{11}$ | $\frac{6}{11}$ | 3 | $-\frac{3}{22}$ | 86° |
| 4 | | | 1 | $-\frac{48}{25}$ | $\frac{36}{25}$ | $-\frac{16}{25}$ | $\frac{3}{25}$ | $\frac{12}{25}$ | 4 | $-\frac{12}{125}$ | 73° |
| 5 | | 1 | $-\frac{300}{137}$ | $\frac{300}{137}$ | $-\frac{200}{137}$ | $\frac{75}{137}$ | $-\frac{12}{137}$ | $\frac{60}{137}$ | 5 | $-\frac{110}{137}$ | 51° |
| 6 | 1 | $-\frac{360}{147}$ | $\frac{450}{147}$ | $-\frac{400}{147}$ | $\frac{225}{147}$ | $-\frac{72}{147}$ | $\frac{10}{147}$ | $\frac{60}{147}$ | 6 | $-\frac{20}{343}$ | 18° |

La partie gauche de la figure 10.10 montre la partie supérieure des régions de stabilité absolue des BDF respectivement à 1 pas (l’extérieur du disque de rayon 1 et de centre 1), à 2 pas et à 3 pas (l’extérieur des régions fermées dans le demi-plan droit). Les angles soutendus à l’origine sont $\alpha = 90^\circ$ dans le deux lers cas et $\alpha = 88^\circ$ dans le 3ème cas. La partie droite de la figure 10.10 montre la partie supérieure des régions de stabilité absolue des BDF respectivement à 4, 5 et 6 pas. Les angles soutendus à l’origine sont respectivement 73° , 51° et 18° . Dans tous ces cas, la région de stabilité absolue inclut l’axe des réels négatifs.

On emploie les BDF pour résoudre des systèmes raides.

10.9.3. Formules de différentiation numérique. Les formules de différentiation numérique (“numerical differentiation formulae”) (NDF) sont une

modification des BDF. Notons

$$\nabla y_n = y_n - y_{n-1}$$

la différence rétrograde de y_n . On récrit les BDF d'ordre $p = k$ à k pas sous la forme

$$\sum_{m=1}^k \frac{1}{m} \nabla^m y_{n+1} = h f_{n+1}.$$

On résout l'équation algébrique par une méthode de Newton simplifiée. On démarre la récurrence avec la valeur prédite

$$y_{n+1}^{[0]} = \sum_{m=0}^k \frac{1}{m} \nabla^m y_n.$$

Puis on continue avec le NDF d'ordre $p = k$ à k pas

$$\sum_{m=1}^k \frac{1}{m} \nabla^m y_{n+1} = h f_{n+1} + \kappa \gamma_k (y_{n+1} - y_{n+1}^{[0]}),$$

où κ est un paramètre scalaire et $\gamma_k = \sum_{j=1}^k 1/j$. Les NDF d'ordre 1 à 5 se trouvent dans le tableau 10.8.

TABLE 10.8. Coefficients des NDF.

| k | κ | α_5 | α_4 | α_3 | α_2 | α_1 | α_0 | β_k | p | C_{p+1} | α |
|---|----------|------------|------------|------------|------------|------------|------------|-----------|-----|-----------|----------|
| 1 | -37/200 | | | | | 1 | -1 | 1 | 1 | 1 | 90° |
| 2 | -1/9 | | | | 1 | -4/3 | 1/3 | 2/3 | 2 | -2/9 | 90° |
| 3 | -0.0823 | | | 1 | -18/11 | 9/11 | -2/11 | 6/11 | 3 | -3/22 | 80° |
| 4 | -0.0415 | | 1 | -48/25 | 36/25 | -16/25 | 3/25 | 12/25 | 4 | -12/125 | 66° |
| 5 | 0 | 1 | -300/137 | 300/137 | -200/137 | 75/137 | -12/137 | 60/137 | 5 | -110/137 | 51° |

Le choix du nombre κ est un compromis pour équilibrer l'efficacité du pas h et de l'angle de stabilité α . Les NDF ont un gain en h de 26% sur les BDF d'ordre $p = 1, 2, 3$, de 12% pour $p = 4$, et 0% pour $p = 5$. Le taux de changement de l'angle de stabilité est respectivement 0%, 0%, -7%, -10%, et 0%. On ne considère aucun NDF d'ordre 6 parce que l'angle α est trop petit.

10.9.4. L'effet d'un grand rapport de raideur. L'exemple suivant analyse l'effet d'un grand rapport de raideur sur le pas pour 5 méthodes de la suite ODE de Matlab appliquées sur un simple système de 2 équations différentielles découplées. On appelle pseudo-raides de tels problèmes puisqu'ils sont résolubles par des méthodes implicites.

Soit le problème à valeurs initiales

$$\begin{bmatrix} y_1(x) \\ y_2(x) \end{bmatrix}' = \begin{bmatrix} -1 & 0 \\ 0 & -10^q \end{bmatrix} \begin{bmatrix} y_1(x) \\ y_2(x) \end{bmatrix}, \quad \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad (10.39)$$

qu'on récrit sous forme matrice-vecteur :

$$\mathbf{y}' = \mathbf{A}\mathbf{y}, \quad \mathbf{y}(0) = \mathbf{y}_0.$$

Puisque les valeurs propres de A sont

$$\lambda_1 = -1, \quad \lambda_2 = -10^q,$$

le rapport de rigidité (10.38) du système est

$$r = 10^q.$$

La solution est

$$\begin{bmatrix} y_1(x) \\ y_2(x) \end{bmatrix} = \begin{bmatrix} e^{-x} \\ e^{-10^q x} \end{bmatrix}.$$

Même si la solution $y_2(x)$ est à décroissance numérique rapide à cause du facteur $\exp(-10^q x)$ pour q grand, le grand rapport de raideur continue de restreindre le pas des méthodes explicites, incluant les paires prédicteur-correcteur.

EXEMPLE 10.13. Étudier l'effet du rapport de raideur sur le nombre de pas utilisés par 5 méthodes de MATLAB pour résoudre le problème (10.39) avec $q = 1$ et $q = 5$.

RÉSOLUTION. La fonction M `exp10_13.m` est

```
function uprime = exp10_13(x,u); % Exemple 10.13
global q % variable globale
A=[-1 0;0 -10^q]; % matrice A
uprime = A*u;
```

Les commandes suivantes résolvent le problème aux valeurs initiales non raide avec $q = 1$ et $r = e^{10}$ et les tolérances relative 10^{-12} et absolue 10^{-14} . Sous l'option `stats on`, le code fait le décompte du nombre d'évaluations des fonctions du 2ème membre.

```
clear;
global q; q=1;
tspan = [0 1]; y0 = [1 1]';
options = odeset('RelTol',1e-12,'AbsTol',1e-14,'Stats','on');
[x23,y23] = ode23('exp10_13',tspan,y0,options);
[x45,y45] = ode45('exp10_13',tspan,y0,options);
[x113,y113] = ode113('exp10_13',tspan,y0,options);
[x23s,y23s] = ode23s('exp10_13',tspan,y0,options);
[x15s,y15s] = ode15s('exp10_13',tspan,y0,options);
```

De même, quand $q = 5$ et $r = \exp(10^5)$ le programme résout le problème aux valeurs initiales pseudo-raide (10.39) avec $q = 1$ et $r = e^{10}$ et les tolérances relative 10^{-12} et absolue 10^{-14} . Le tableau 10.9 liste le nombre de pas utilisés avec $q = 1$ et $q = 5$ par 5 méthodes de la suite ODE.

On voit, d'après le tableau, que les solveurs explicites sont très lents et très coûteux sur des équations pseudo-raides. \square

On considère un autre exemple d'équation du second ordre qui contient un paramètre q . On trouve d'abord la solution analytique.

EXEMPLE 10.14. Résoudre le système

$$y'' + (10^q + 1)y' + 10^q y = 0 \quad \text{sur } [0, 1],$$

aux valeurs initiales

$$y(0) = 2, \quad y'(0) = -10^q - 1,$$

TABLE 10.9. Nombre de pas utilisés avec $q = 1$ et $q = 5$ et tolérances relative et absolue $RT = 10^{-3}$ et $AT = 10^{-6}$ par défaut et tolérances $RT = 10^{-12}$ et $AT = 10^{-14}$.

| (RT, AT) | $(10^{-3}, 10^{-6})$ | | $(10^{-12}, 10^{-14})$ | |
|------------|----------------------|--------|------------------------|--------|
| | 1 | 5 | 1 | 5 |
| ode23 | 29 | 39 823 | 24 450 | 65 944 |
| ode45 | 13 | 30 143 | 601 | 30 856 |
| ode113 | 28 | 62 371 | 132 | 64 317 |
| ode23s | 37 | 57 | 30 500 | 36 925 |
| ode15s | 43 | 89 | 773 | 1 128 |

avec un paramètre réel q .

RÉSOLUTION. On substitue

$$y(x) = e^{\lambda x}$$

dans l'équation différentielle pour obtenir le polynôme caractéristique et les valeurs propres :

$$\lambda^2 + (10^q + 1)\lambda + 10^q = (\lambda + 10^q)(\lambda + 1) = 0 \implies \lambda_1 = -10^q, \quad \lambda_2 = -1.$$

On a les 2 solutions indépendantes

$$y_1 = e^{-10^q x}, \quad y_2(x) = e^{-x}$$

et la solution générale

$$y(x) = c_1 e^{-10^q x} + c_2 e^{-x}.$$

On obtient $c_1 = 1$ et $c_2 = 1$ au moyen des conditions initiales. Donc la solution unique est

$$y(x) = e^{-10^q x} + e^{-x}. \quad \square$$

On transforme le problème de l'exemple 10.14 en un système de 2 équations couplées du 1er ordre pour trouver la solution par Matlab numérique.

EXEMPLE 10.15. Reformuler le problème

$$y'' + (10^q + 1)y' + 10^q y = 0 \quad \text{sur } [0, 1],$$

aux conditions initiales

$$y(0) = 2, \quad y'(0) = -10^q - 1,$$

avec paramètre réel q , en un système de 2 équations du 1er ordre et trouver le vecteur solution.

RÉSOLUTION. Posons

$$u_1 = y, \quad u_2 = y'.$$

Alors,

$$u_2 = u_1', \quad u_2' = y'' = -10^q u_1 - (10^q + 1)u_2.$$

On obtient donc le système $\mathbf{u}' = \mathbf{A}\mathbf{u}$ aux valeurs initiales :

$$\begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -10^q & -(10^q + 1) \end{bmatrix} \begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix}, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} = \begin{bmatrix} 2 \\ -10^q - 1 \end{bmatrix}.$$

On substitue la fonction vectorielle

$$\mathbf{u}(x) = \mathbf{c} e^{\lambda x}$$

dans le système différentiel pour obtenir le problème aux valeurs propres

$$(A - \lambda I)\mathbf{c} = \begin{bmatrix} -\lambda & 1 \\ -10^q & -(10^q + 1) - \lambda \end{bmatrix} \mathbf{c} = 0,$$

Ce problème admet une solution non nulle \mathbf{c} si et seulement si

$$\det(A - \lambda I) = \lambda^2 + (10^q + 1)\lambda + 10^q = (\lambda + 10^q)(\lambda + 1) = 0.$$

On a donc les valeurs propres

$$\lambda_1 = -10^q, \quad \lambda_2 = -1.$$

On trouve les vecteurs propres solutions des systèmes linéaires

$$(A - \lambda_i I)\mathbf{v}_i = 0.$$

Alors,

$$\begin{bmatrix} 10^q & 1 \\ -10^q & -1 \end{bmatrix} \mathbf{v}_1 = 0 \implies \mathbf{v}_1 = \begin{bmatrix} 1 \\ -10^q \end{bmatrix}$$

et

$$\begin{bmatrix} 1 & 1 \\ -10^q & -10^q \end{bmatrix} \mathbf{v}_2 = 0 \implies \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

On a la solution générale

$$\mathbf{u}(x) = c_1 e^{-10^q x} \mathbf{v}_1 + c_2 e^{-x} \mathbf{v}_2.$$

On obtient $c_1 = 1$ et $c_2 = 1$ au moyen des conditions initiales. Donc la solution unique est

$$\begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix} = \begin{bmatrix} 1 \\ -10^q \end{bmatrix} e^{-10^q x} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-x}. \quad \square$$

On voit que le rapport de raideur de l'équation de l'exemple 10.15 est

$$10^q.$$

EXEMPLE 10.16. Résoudre l'équation différentielle non raide avec $q = 1$:

$$y'' + (10^q + 1)y' + 10^q y = 0 \quad \text{sur } [0, 1],$$

aux conditions initiales

$$y(0) = 2, \quad y'(0) = -10^q - 1,$$

par 5 solveurs Matlab et comparer le nombre de pas employés.

RÉSOLUTION. La fonction M `exp10_16.m` est

```
function uprime = exp10_16(x,u)
global q
A=[0 1;-10^q -1-10^q];
uprime = A*u;
```

Les commandes suivantes résolvent le problème.

```

>> clear
>> global q; q = 1;
>> xspan = [0 1]; u0 = [2 -(10^q + 1)]';
>> [x23,u23] = ode23('exp10_16',xspan,u0);
>> [x45,u45] = ode45('exp10_16',xspan,u0);
>> [x113,u113] = ode113('exp10_16',xspan,u0);
>> [x23s,u23s] = ode23s('exp10_16',xspan,u0);
>> [x15s,u15s] = ode15s('exp10_16',xspan,u0);
>> whos

```

| Name | Size | Bytes | Class |
|-------|------|-------|-----------------------|
| q | 1x1 | 8 | double array (global) |
| u0 | 2x1 | 16 | double array |
| u113 | 26x2 | 416 | double array |
| u15s | 32x2 | 512 | double array |
| u23 | 20x2 | 320 | double array |
| u23s | 25x2 | 400 | double array |
| u45 | 49x2 | 784 | double array |
| x113 | 26x1 | 208 | double array |
| x15s | 32x1 | 256 | double array |
| x23 | 20x1 | 160 | double array |
| x23s | 25x1 | 200 | double array |
| x45 | 49x1 | 392 | double array |
| xspan | 1x2 | 16 | double array |

Grand total is 461 elements using 3688 bytes

On voit dans le tableau produit par la commande `whos` que les solveurs non raides `ode23`, `ode45`, `ode113`, et les solveurs raides `ode23s`, `ode15s`, emploient respectivement 20, 49, 26, et 25, 32 pas. \square

EXEMPLE 10.17. Résoudre l'équation différentielle pseudo-raide avec $q = 5$:

$$y'' + (10^q + 1)y' + 10^q y = 0 \quad \text{sur } [0, 1],$$

aux conditions initiales

$$y(0) = 2, \quad y'(0) = -10^q - 1,$$

par 5 solveurs Matlab et comparer le nombre de pas employés.

RÉSOLUTION. On met $q = 5$ dans le programme de l'exemple 10.16. On a les résultats suivants.

```

clear
global q; q = 5;
xspan = [0 1]; u0 = [2 -(10^q + 1)]';
[x23,u23] = ode23('exp10_16',xspan,u0);
[x45,u45] = ode45('exp10_16',xspan,u0);
[x113,u113] = ode113('exp10_16',xspan,u0);
[x23s,u23s] = ode23s('exp10_16',xspan,u0);
[x15s,u15s] = ode15s('exp10_16',xspan,u0);

```

La commande `whos` produit le tableau suivant.

| Name | Size | Bytes | Class |
|-------|----------|---------|-----------------------|
| q | 1x1 | 8 | double array (global) |
| u0 | 2x1 | 16 | double array |
| u113 | 62258x2 | 996128 | double array |
| u15s | 107x2 | 1712 | double array |
| u23 | 39834x2 | 637344 | double array |
| u23s | 75x2 | 1200 | double array |
| u45 | 120593x2 | 1929488 | double array |
| x113 | 62258x1 | 498064 | double array |
| x15s | 107x1 | 856 | double array |
| x23 | 39834x1 | 318672 | double array |
| x23s | 75x1 | 600 | double array |
| x45 | 120593x1 | 964744 | double array |
| xspan | 1x2 | 16 | double array |

Grand total is 668606 elements using 5348848 bytes

On voit dans le tableau produit par la commande `whos` que les solveurs non raides `ode23`, `ode45`, `ode113`, et les solveurs raides `ode23s`, `ode15s`, emploient respectivement 39 834, 120 593, 62 258 et 75, 107 pas. Il suit que les solveurs non raides sont très lents pour résoudre des équations raides. \square

On trouve dans MATLAB numérique 4 solveurs avec interpolant gratuit pour les systèmes raides. L'ordre des trois lers codes est bas.

- `ode23s` implémente une modification d'une paire de Rosenbrock (2,3) sans extrapolation locale. avec jacobiens calculés numériquement par défaut.
- `ode23t` implémente la méthode des trapèzes.
- `ode23tb` implémente une méthode de Runge–Kutta implicite à 2 étages.
- Le code `ode15s` est à pas et ordre variables. Le pas est quasi-constant. Il utilise les différences rétrogrades d'une famille de formules de Klopfenstein–Shampine appelées “Numerical Differentiation Formulae” d'ordre 1 à 5. sans extrapolation locale et avec jacobiens calculés numériquement par défaut. En option, on peut utiliser les BDF de Gear.

On trouve les détails sur ces méthodes dans la référence *The MATLAB ODE Suite*, L. F. Shampine et M. W. Reichelt, SIAM Journal on Scientific Computing, **18**(1), 1997.

Enfin, le code `ode15i` est une méthode à ordre variable pour résoudre de systèmes complètement implicites

Calcul matriciel

Le calcul matriciel occupe une place centrale dans les logiciels modernes d'informatique depuis la venue des ordinateurs en sciences et en ingénierie. Dans le présent chapitre, on étudie les algorithmes pour la résolution de systèmes linéaires :

$$A\mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{m \times n},$$

et de problèmes de valeurs propres :

$$A\mathbf{x} = \lambda\mathbf{x}, \quad A \in \mathbb{R}^{n \times n}, \quad \mathbf{x} \neq 0,$$

tels que mis en œuvres dans les logiciels modernes où la précision, la stabilité et la complexité sont très importantes

11.1. Solution LU de $A\mathbf{x} = \mathbf{b}$

On obtient la solution d'un système linéaire

$$A\mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{n \times n},$$

avec pivotage partiel, expliqué ci-après, par la décomposition LU de A ,

$$A = LU,$$

où L provient d'une permutation des lignes d'une matrice triangulaire inférieure M avec $m_{ii} = 1$ et $|m_{ij}| \leq 1$, for $i > j$, et U est une matrice triangulaire supérieure. On obtient donc le système

$$LU\mathbf{x} = \mathbf{b}.$$

La solution se fait en deux étapes. D'abord on résout

$$L\mathbf{y} = \mathbf{b}$$

pour \mathbf{y} par substitution prograde (avant) ; ensuite on résout

$$U\mathbf{x} = \mathbf{y}$$

pour \mathbf{x} par substitution rétrograde (arrière). L'exemple suivant illustre ces deux étapes.

EXEMPLE 11.1. Résoudre le système $A\mathbf{x} = \mathbf{b}$,

$$\begin{bmatrix} 3 & 9 & 6 \\ 18 & 48 & 39 \\ 9 & -27 & 42 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 23 \\ 136 \\ 45 \end{bmatrix}$$

par décomposition LU avec **pivotage partiel**.

RÉSOLUTION. Le pivot maximum en module dans la 1ère colonne de A est $a_{21} = 18$ puisque

$$|18| > |3|, \quad |18| > |9|.$$

On permute donc la 2ème et la 1ère lignes de A :

$$P_1 A = \begin{bmatrix} 18 & 48 & 39 \\ 3 & 9 & 6 \\ 9 & -27 & 42 \end{bmatrix}, \quad \text{où } P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ensuite on applique une transformation gaussienne, M_1 , sur $P_1 A$ avec multiplicateurs $-1/6$ et $-1/2$ pour annuler les éléments sous le 18 dans la 1ère colonne :

$$\begin{bmatrix} 1 & 0 & 0 \\ -1/6 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 18 & 48 & 39 \\ 3 & 9 & 6 \\ 9 & -27 & 42 \end{bmatrix} = \begin{bmatrix} 18 & 48 & 39 \\ 0 & 1 & -1/2 \\ 0 & -51 & 45/2 \end{bmatrix}.$$

Donc

$$M_1 P_1 A = A_1.$$

Dans la sous-matrice d'ordre 2 :

$$\begin{bmatrix} 1 & -1/2 \\ -51 & 45/2 \end{bmatrix},$$

le pivot de la 1ère colonne est -51 puisque

$$|-51| > |1|.$$

Donc on permute la 2ème et la 3ème lignes :

$$P_2 A_1 = \begin{bmatrix} 18 & 48 & 39 \\ 0 & -51 & 45/2 \\ 0 & 1 & -1/2 \end{bmatrix}, \quad \text{où } P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Pour annuler l'élément $(3,2)$ on applique une transformation gaussienne M_2 sur $P_2 A_1$ avec multiplicateurs $1/51$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/51 & 1 \end{bmatrix} \begin{bmatrix} 18 & 48 & 39 \\ 0 & -51 & 45/2 \\ 0 & 1 & -1/2 \end{bmatrix} = \begin{bmatrix} 18 & 48 & 39 \\ 0 & -51 & 22.5 \\ 0 & 0 & -0.0588 \end{bmatrix}.$$

Alors

$$M_2 P_2 A_1 = U.$$

Donc

$$M_2 P_2 M_1 P_1 A = U,$$

et

$$A = P_1^{-1} M_1^{-1} P_2^{-1} M_2^{-1} U = LU.$$

Les transformations gaussiennes inverses s'obtiennent facilement :

$$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ -b & 0 & 1 \end{bmatrix} \implies M_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{bmatrix},$$

$$M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -c & 1 \end{bmatrix} \implies M_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix},$$

si l'on connaît les multiplicateurs $-a$, $-b$, $-c$. De plus, le produit $M_1^{-1}M_2^{-1}$ est tout aussi simple :

$$M_1^{-1}M_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{bmatrix}.$$

On sait qu'une permutation P , représentée par la matrice identité I avec lignes permutées, est une matrice orthogonale. Alors,

$$P^{-1} = P^T.$$

Donc, si

$$L = P_1^T M_1^{-1} P_2^T M_2^{-1},$$

par simple réarrangement des éléments de M_1^{-1} et M_2^{-1} sans opérations arithmétiques, on obtient la matrice

$$\begin{aligned} L &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1/6 & 1 & 0 \\ 1/2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/51 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1/6 & 1 & 0 \\ 1 & 0 & 0 \\ 1/2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/51 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1/6 & -1/51 & 1 \\ 1 & 0 & 0 \\ 1/2 & 1 & 0 \end{bmatrix} \end{aligned}$$

qui provient d'une permutation des lignes d'une matrice triangulaire inférieure avec 1 sur la diagonale, c'est-à-dire, L devient triangulaire inférieure avec 1 sur la diagonale si l'on permute les 2 premières lignes puis l'on permute la nouvelle 2ème ligne et la 3ème ligne : $P_2 P_1 L$ est triangulaire inférieure.

On résout le système

$$L\mathbf{y} = \mathbf{b}$$

par substitution prograde:

$$\begin{bmatrix} 1/6 & -1/51 & 1 \\ 1 & 0 & 0 \\ 1/2 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 23 \\ 136 \\ 45 \end{bmatrix},$$

$$y_1 = 136,$$

$$y_2 = 45 - 136/2 = -23,$$

$$y_3 = 23 - 136/6 - 23/51 = -0.1176.$$

Puis on résout le système

$$U\mathbf{x} = \mathbf{y}$$

par substitution rétrograde:

$$\begin{bmatrix} 18 & 48 & 39 \\ 0 & -51 & 22.5 \\ 0 & 0 & -0.0588 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 136 \\ -23 \\ -0.1176 \end{bmatrix},$$

$$x_3 = 0.1176/0.0588 = 2,$$

$$x_2 = (-23 - 22.5 \times 2)/(-51) = 1.3333,$$

$$x_1 = (136 - 48 \times 1.3333 - 39 \times 2)/18 = -0.3333.$$

□

La session de Matlab qui suit fait exactement la même chose.

```
>> A = [3 9 6; 18 48 39; 9 -27 42]
```

```
A =  
    3     9     6  
   18    48    39  
    9   -27    42
```

```
>> [L,U] = lu(A)
```

```
L =  
    0.1667   -0.0196    1.0000  
    1.0000         0         0  
    0.5000    1.0000         0
```

```
U =  
   18.0000   48.0000   39.0000  
         0  -51.0000   22.5000  
         0         0   -0.0588
```

```
>> b = [23; 136; 45]
```

```
b =  
    23  
   136  
    45
```

```
>> y = L\b % substitution prograde
```

```
y =  
   136.0000  
  -23.0000  
   -0.1176
```

```
>> x = U\y % substitution retrograde
```

```
x =  
  -0.3333  
   1.3333  
   2.0000
```

```
>> z = A\b % l'inverse a gauche de Matlab resout Az = b par decomposition LU
```

```
z =  
  -0.3333  
   1.3333  
   2.0000
```

La commande pédagogique de Matlab :

$[L,U,P] = \text{lu}(A)$

trouve la matrice de permutation P qui fait d'un seul coup tout le pivotage du système

$$Ax = b,$$

produit le système permuté équivalent :

$$PAx = Pb$$

et calcule la décomposition LU de PA :

$$PA = LU,$$

où la matrice L triangulaire inférieure avec 1 sur la diagonale, $l_{ij} \leq 1$, pour $i > j$, et la matrice U est triangulaire supérieure.

On répète la session de Matlab avec la matrice P .

$A = [3 \ 9 \ 6; \ 18 \ 48 \ 39; \ 9 \ -27 \ 42]$

```
A =
     3     9     6
    18    48    39
     9   -27    42
```

$b = [23; \ 136; \ 45]$

```
b =
    23
   136
    45
```

$[L,U,P] = \text{lu}(A)$

```
L =
   1.0000         0         0
   0.5000    1.0000         0
   0.1667   -0.0196    1.0000
```

```
U = 18.0000   48.0000   39.0000
      0  -51.0000   22.5000
      0         0   -0.0588
```

```
P =
     0     1     0
     0     0     1
     1     0     0
```

$y = L \setminus P * b$

```
y = 136.0000
    -23.0000
    -0.1176
```

$x = U \setminus y$

```
x = -0.3333
     1.3333
     2.0000
```

THÉORÈME 11.1. *La décomposition LU d'une matrice A existe si et seulement si tous les mineurs principaux de A sont nonnuls.*

Les mineurs principaux de A sont les déterminants supérieurs gauches des sous-matrices de A . Le pivotage partiel essaie d'obtenir des mineurs principaux de PA nonnuls.

EXEMPLE 11.2. Soit A et b :

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 12 & 13 & 6 \\ -3 & 8 & 9 \end{bmatrix}, \quad b = \begin{bmatrix} 14 \\ 40 \\ -28 \end{bmatrix}.$$

Décomposer A en LU **sans pivotage** et résoudre le système :

$$Ax = b.$$

RÉSOLUTION. Pour $M_1A = A_1$, on a :

$$\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 \\ 12 & 13 & 6 \\ -3 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 5 & 6 \\ 0 & 10 & 9 \end{bmatrix}.$$

Pour $M_2A_1 = U$, on a

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 \\ 0 & 5 & 6 \\ 0 & 10 & 9 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 5 & 6 \\ 0 & 0 & -3 \end{bmatrix} = U,$$

c'est-à-dire

$$M_2M_1A = U, \quad A = M_1^{-1}M_2^{-1}U = LU.$$

Alors

$$L = M_1^{-1}M_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}.$$

On obtient \mathbf{y} pour $L\mathbf{y} = \mathbf{b}$ par substitution prograde :

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 14 \\ 40 \\ -28 \end{bmatrix};$$

alors

$$\begin{aligned} y_1 &= 14, \\ y_2 &= 40 - 56 = -16, \\ y_3 &= -28 + 14 + 32 = 18. \end{aligned}$$

Enfin, on obtient \mathbf{x} pour $U\mathbf{x} = \mathbf{y}$ par substitution rétrograde :

$$\begin{bmatrix} 3 & 2 & 0 \\ 0 & 5 & 6 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 14 \\ -16 \\ 18 \end{bmatrix};$$

alors

$$\begin{aligned} x_3 &= -6, \\ x_2 &= (-16 + 36)/5 = 4, \\ x_1 &= (14 - 8)/3 = 2. \end{aligned}$$

□

On remarque que sans pivotage $|l_{ij}|$, $i > j$, peut être supérieur à 1.

La décomposition LU sans pivotage partiel est instable et peut causer des erreurs importantes dans la solution. En pratique, le pivotage partiel est ordinairement stable. Cependant, dans certains cas, on doit recourir au pivotage complet sur les lignes et les colonnes pour des raisons de stabilité. On peut aussi employer la décomposition stable QR.

Parfois il est utile d'étalonner les lignes ou les colonnes de la matrice d'un système linéaire avant de le résoudre. Ceci peut modifier le choix des pivots. En pratique, il faut considérer la signification et la dimension physique des inconnues pour bien étalonner ou bien équilibrer la matrice. Certains logiciels offrent ces options. À l'exemple suivant on étalonne les lignes dans la norme l_∞ .

EXEMPLE 11.3. Soit le système :

$$\begin{aligned} 30.00x_1 + 591400x_2 &= 591700 \\ 5.29x_1 - 6.130x_2 &= 46.70 \end{aligned}$$

Étalonner chaque équation en norme l_∞ pour que le plus grand coefficient de chaque ligne du 1er membre soit au plus 1 en module et résoudre par décomposition LU avec pivotage en arithmétique à 4 décimales.

RÉSOLUTION. On divise la 1ère équation par

$$s_1 = \max\{|30.00|, |591400|\} = 591400$$

et la seconde par

$$s_2 = \max\{|5.291|, |6.130|\} = 6.130.$$

Alors, on obtient

$$\frac{|a_{11}|}{s_1} = \frac{30.00}{591400} = 0.5073 \times 10^{-4}, \quad \frac{|a_{21}|}{s_2} = \frac{5.291}{6.130} = 0.8631.$$

Donc le pivot étalonné est dans la seconde équation. Pour éviter les erreurs d'arrondi de la division, on applique la décomposition LU sur le système original avec le pivot trouvé plutôt que sur le système étalonné :

$$\begin{aligned} 5.29x_1 - 6.130x_2 &= 46.70 \\ 30.00x_1 + 591400x_2 &= 591700. \end{aligned}$$

On obtient alors le résultat correct:

$$x_1 = 10.00, \quad x_2 = 1.000.$$

D'autre part, la décomposition LU en arithmétique à 4 décimales sans pivotage produit le résultat erroné $x_1 \approx -10.00$ et $x_2 \approx 1.001$. \square

On trouve sur le réseau au lien <ftp://ftp.cs.cornell.edu/pub/cv> une fonction M en Matlab qui résout un système linéaire par décomposition LU. L'algorithme substitution prograde résout un système triangulaire inférieur :

```
function x = LTriSol(L,b)
%
% Pre:
%   L   n-by-n nonsingular lower triangular matrix
%   b   n-by-1
%
% Post:
```

```

% x Lx = b

n = length(b);
x = zeros(n,1);
for j=1:n-1
    x(j) = b(j)/L(j,j);
    b(j+1:n) = b(j+1:n) - L(j+1:n,j)*x(j);
end
x(n) = b(n)/L(n,n);

```

L'algorithme substitution rétrograde résout un système triangulaire supérieur :

```

function x = UTriSol(U,b)
%
% Pre:
% U n-by-n nonsingular upper triangular matrix
% b n-by-1
%
% Post:
% x Lx = b

n = length(b);
x = zeros(n,1);
for j=n:-1:2
    x(j) = b(j)/U(j,j);
    b(1:j-1) = b(1:j-1) - x(j)*U(1:j-1,j);
end
x(1) = b(1)/U(1,1);

```

La fonction suivante fait la décomposition LU sans pivotage :

```

function [L,U] = GE(A);
%
% Pre:
% A n-by-n
%
% Post:
% L n-by-n unit lower triangular with |L(i,j)| <= 1.
% U n-by-n upper triangular.
% A = LU

```

```

[n,n] = size(A);
for k=1:n-1
    A(k+1:n,k) = A(k+1:n,k)/A(k,k);
    A(k+1:n,k+1:n) = A(k+1:n,k+1:n) - A(k+1:n,k)*A(k,k+1:n);
end
L = eye(n,n) + tril(A,-1);
U = triu(A);

```

La fonction suivante fait la décomposition LU avec pivotage partiel:

```

function [L,U,piv] = GEpiv(A);
%

```

```

% Pre:
% A      n-by-n
%
% Post:
% L      n-by-n unit lower triangular with |L(i,j)| <= 1.
% U      n-by-n upper triangular
% piv    integer n-vecteur that is a permutation of 1:n.
%
%        A(piv,:) = LU

[n,n] = size(A);
    piv = 1:n;
for k=1:n-1
    [maxv,r] = max(abs(A(k:n,k)));
    q = r+k-1;
    piv([k q]) = piv([q k]);
    A([k q],:) = A([q k],:);
    if A(k,k) ~= 0
        A(k+1:n,k) = A(k+1:n,k)/A(k,k);
        A(k+1:n,k+1:n) = A(k+1:n,k+1:n) - A(k+1:n,k)*A(k,k+1:n);
    end
end
    L = eye(n,n) + tril(A,-1);
    U = triu(A);

```

11.2. La décomposition de Cholesky

L'importante classe de matrices définies positives admet la décomposition de Cholesky

$$A = GG^T$$

où G est triangulaire inférieure.

DÉFINITION 11.1. Une matrice symétrique $A \in \mathbb{R}^{n \times n}$ est définie positive si

$$\mathbf{x}^T A \mathbf{x} > 0, \quad \text{pour tout } \mathbf{x} \neq 0, \quad \mathbf{x} \in \mathbb{R}^n.$$

Dans ce cas on écrit $A > 0$.

Une matrice symétrique est définie positive si et seulement si toutes ses valeurs propres λ ,

$$A\mathbf{x} = \lambda\mathbf{x}, \quad \mathbf{x} \neq 0,$$

sont positives, $\lambda > 0$.

Une matrice symétrique est définie positive si et seulement si tous ses mineurs principaux sont positifs. Par exemple,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} > 0$$

si et seulement si

$$\det a_{11} = a_{11} > 0, \quad \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} > 0, \quad \det A > 0.$$

Si $A > 0$, alors $a_{ii} > 0$, $i = 1, 2, \dots, n$.

Une matrice A d'ordre n est à *diagonale dominante* si

$$|a_{ii}| > |a_{i1}| + |a_{i2}| + \dots + |a_{i,i-1}| + |a_{i,i+1}| + \dots + |a_{in}|, \quad i = 1, 2, \dots, n.$$

Une matrice symétrique à diagonale dominante et à diagonale positive est définie positive.

THÉORÈME 11.2. *Soit A définie positive. Alors la décomposition de Cholesky*

$$A = GG^T$$

peut se faire sans pivotage. Donc on peut résoudre $A\mathbf{x} = \mathbf{b}$ par substitutions prograde et rétrograde:

$$G\mathbf{y} = \mathbf{b}, \quad G^T\mathbf{x} = \mathbf{y}.$$

EXEMPLE 11.4. Soit

$$A = \begin{bmatrix} 4 & 6 & 8 \\ 6 & 34 & 52 \\ 8 & 52 & 129 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ -160 \\ -452 \end{bmatrix}.$$

Faire la décomposition de Cholesky de A et employer cette décomposition pour calculer le déterminant de A et résoudre le système

$$A\mathbf{x} = \mathbf{b}.$$

RÉSOLUTION. On obtient la décomposition de Cholesky (sans pivotage) en résolvant le système suivant pour g_{ij} :

$$\begin{bmatrix} g_{11} & 0 & 0 \\ g_{21} & g_{22} & 0 \\ g_{31} & g_{32} & g_{33} \end{bmatrix} \begin{bmatrix} g_{11} & g_{21} & g_{31} \\ 0 & g_{22} & g_{32} \\ 0 & 0 & g_{33} \end{bmatrix} = \begin{bmatrix} 4 & 6 & 8 \\ 6 & 34 & 52 \\ 8 & 52 & 129 \end{bmatrix}.$$

$$\begin{aligned} g_{11}^2 &= 4 && \implies g_{11} = 2 > 0, \\ g_{11}g_{21} &= 6 && \implies g_{21} = 3, \\ g_{11}g_{31} &= 8 && \implies g_{31} = 4, \\ g_{21}^2 + g_{22}^2 &= 34 && \implies g_{22} = 5 > 0, \\ g_{21}g_{31} + g_{22}g_{32} &= 52 && \implies g_{32} = 8, \\ g_{31}^2 + g_{32}^2 + g_{33}^2 &= 129 && \implies g_{33} = 7 > 0. \end{aligned}$$

Donc

$$G = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 5 & 0 \\ 4 & 8 & 7 \end{bmatrix},$$

et

$$\det A = \det G \det G^T = (\det G)^2 = (2 \times 5 \times 7)^2 > 0.$$

On résout $G\mathbf{y} = \mathbf{b}$ par substitution prograde :

$$\begin{bmatrix} 2 & 0 & 0 \\ 3 & 5 & 0 \\ 4 & 8 & 7 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -160 \\ -452 \end{bmatrix},$$

d'où

$$\begin{aligned}y_1 &= 0, \\y_2 &= -32, \\y_3 &= (-452 + 256)/7 = -28.\end{aligned}$$

On résout $G^T \mathbf{x} = \mathbf{y}$ par substitution rétrograde :

$$\begin{bmatrix} 2 & 3 & 4 \\ 0 & 5 & 8 \\ 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -32 \\ -28 \end{bmatrix},$$

d'où

$$\begin{aligned}x_3 &= -4, \\x_2 &= (-32 + 32)/5 = 0, \\x_1 &= (0 - 3 \times 0 + 16)/2 = 8. \quad \square\end{aligned}$$

La commande `chol` de Matlabf numérique fait la décompositon de Cholesky $R^T R$ d'une matrice A définie positive :

```
>> A = [4 6 8;6 34 52;8 52 129];
>> R = chol(A)
```

R =

```
 2     3     4
 0     5     8
 0     0     7
```

On trouve les M fonctions Matlab suivantes au lien <ftp://ftp.cs.cornell.edu/pub/cv>. Ces fonctions illustrent les différents niveaux de multiplication matrice-vecteur.

On a la décompositon de Cholesky "scalaire" la plus simple :

```
function G = CholScalar(A);
%
% Pre: A is a symmetric et positive definite matrix.
% Post: G is lower triangular et A = G*G'.

[n,n] = size(A);
G = zeros(n,n);
for i=1:n
    % Compute G(i,1:i)
    for j=1:i
        s = A(j,i);
        for k=1:j-1
            s = s - G(j,k)*G(i,k);
        end
        if j<i
            G(i,j) = s/G(j,j);
        else
            G(i,i) = sqrt(s);
        end
    end
end
```

end

Le produit scalaire de 2 vecteurs, $c = x^T y$, est un scalaire. Puisque la k ème boucle de `CholScalar` gère le produit scalaire de sous-lignes de G , on obtient l'implémentation du produit scalaire de niveau 1.

```
function G = CholDot(A);
%
% Pre: A is a symmetric et positive definite matrix.
% Post: G is lower triangular et A = G*G'.

[n,n] = size(A);
G = zeros(n,n);
for i=1:n
    % Compute G(i,1:i)
    for j=1:i
        if j==1
            s = A(j,i);
        else
            s = A(j,i) - G(j,1:j-1)*G(i,1:j-1)';
        end
        if j<i
            G(i,j) = s/G(j,j);
        else
            G(i,i) = sqrt(s);
        end
    end
end
```

Une mise à jour de la forme

$$\text{vecteur} \leftarrow \text{vecteur} + \text{vecteur} \cdot \text{scalaire}$$

est un opération *saxpy*, qui est l'anachronyme de "scalaire a fois x plus y ", c'est-à-dire $y = ax + y$.

On a une implémentation en version colonne qui emploie les opérations saxpy :

```
function G = CholSax(A);
%
% Pre: A is a symmetric et positive definite matrix.
% Post: G is lower triangular et A = G*G'.

[n,n] = size(A);
G = zeros(n,n);
s = zeros(n,1);
for j=1:n
    s(j:n) = A(j:n,j);
    for k=1:j-1
        s(j:n) = s(j:n) - G(j:n,k)*G(j,k);
    end
    G(j:n,j) = s(j:n)/sqrt(s(j));
end
```

Une mise à jour de la forme

$$\text{vecteur} \leftarrow \text{vecteur} + \text{matrice} \times \text{vecteur}$$

est un opération *gaxpy*, qui est l’anachronyme de “générale A fois x plus y ” (“général saxpy”), c’est-à-dire $y = Ax + y$. On a une version qui utilise les opérations *gaxpy* de niveau 2.

```
function G = CholGax(A);
%
% Pre: A is a symmetric et positive definite matrix.
% Post: G is lower triangular et A = G*G'.

[n,n] = size(A);
G = zeros(n,n);
s = zeros(n,1);
for j=1:n
    if j==1
        s(j:n) = A(j:n,j);
    else
        s(j:n) = A(j:n,j) - G(j:n,1:j-1)*G(j,1:j-1)';
    end
    G(j:n,j) = s(j:n)/sqrt(s(j));
end
```

Il y a aussi une implémentation récursive qui calcule le facteur de Cholesky ligne par ligne, tout comme *ChoScalar* :

```
function G = CholRecur(A);
%
% Pre: A is a symmetric et positive definite matrix.
% Post: G is lower triangular et A = G*G'.

[n,n] = size(A);
if n==1
    G = sqrt(A);
else
    G(1:n-1,1:n-1) = CholRecur(A(1:n-1,1:n-1));
    G(n,1:n-1) = LTriSol(G(1:n-1,1:n-1),A(1:n-1,n))';
    G(n,n) = sqrt(A(n,n) - G(n,1:n-1)*G(n,1:n-1)');
end
```

Il y a même une implémentation *CholBlock* haute performance de niveau 3 de la décomposition de Cholesky.

11.3. Normes matricielles

En calcul matriciel on emploie des normes pour quantifier des résultats, tels l’estimation de l’erreur et l’étude de la convergence de schémas récurifs..

Soit une matrice $A \in \mathbb{R}^{n \times n}$ ou $\mathbb{C}^{n \times n}$, et un vecteur $\mathbf{x} \in \mathbb{R}^n$ ou \mathbb{C}^n de norme $\|\mathbf{x}\|$. Une *norme matricielle subordonnée* $\|A\|$ est définie par le supremum

$$\|A\| = \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|.$$

En calcul scientifique, il y a 3 normes vectorielles importantes : la *norme* l_1 d'un vecteur \mathbf{x} ,

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \cdots + |x_n|,$$

la *norme euclidienne*, ou *norme* l_2 , de \mathbf{x} ,

$$\|\mathbf{x}\|_2 = \left[\sum_{i=1}^n |x_i|^2 \right]^{1/2} = [|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2]^{1/2},$$

et la *norme supremum* ou uniforme, ou *norme* l_∞ de \mathbf{x} ,

$$\|\mathbf{x}\|_\infty = \sup_{i=1,2,\dots,n} |x_i| = \sup\{|x_1|, |x_2|, \dots, |x_n|\}.$$

On peut démontrer les formules suivantes.

La *norme* l_1 ou colonne maximum de A :

$$\|A\|_1 = \max_{j=1,2,\dots,n} \sum_{i=1}^n |a_{ij}| \quad (\text{colonne maximum dans la norme vectorielle } l_1),$$

la *norme* l_∞ ou ligne maximum de A :

$$\|A\|_\infty = \max_{i=1,2,\dots,n} \sum_{j=1}^n |a_{ij}| \quad (\text{ligne maximum dans la norme vectorielle } l_1),$$

et la *norme* l_2 de A :

$$\|A\|_2 = \max_{i=1,2,\dots,n} \{\sigma_i\} \quad (\text{valeur singulière maximum de } A),$$

où les $\sigma_i^2 \geq 0$ sont les valeurs propres de $A^T A$. On considérera les valeurs singulières d'une matrice à la sous-section 11.9.

La *norme de Frobenius*, ou *norme matricielle euclidienne*,

$$\|A\|_F = \left[\sum_{j=1}^n \sum_{i=1}^n |a_{ij}|^2 \right]^{1/2},$$

est une importante norme matricielle non-subordonnée.

DÉFINITION 11.2 (Conditionnement). Le *conditionnement* d'une matrice $A \in \mathbb{R}^{n \times n}$ est le nombre

$$\kappa(A) = \|A\| \|A^{-1}\|. \quad (11.1)$$

Remarquer que $\kappa(A) \geq 1$ si $\|I\| = 1$.

Le conditionnement de A entre dans la borne supérieure de l'erreur relative de la solution d'un système linéaire

$$A\mathbf{x} = \mathbf{b}.$$

De fait, soit $\hat{\mathbf{x}}$ le solution exacte du système perturbé

$$(A + \Delta A)\hat{\mathbf{x}} = \mathbf{b} + \delta\mathbf{b},$$

où toutes les erreurs expérimentales et d'arrondis se trouvent dans ΔA et $\delta \mathbf{b}$. Alors, l'erreur relative satisfait l'inégalité :

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \left[\frac{\|\Delta A\|}{\|A\|} + \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|} \right]. \quad (11.2)$$

Le système $A\mathbf{x} = \mathbf{b}$ est bien conditionné si $\kappa(A)$ est petit; sinon il est mal conditionné.

EXEMPLE 11.5. Étudier le mauvais conditionnement du système

$$\begin{bmatrix} 1.0001 & 1 \\ 1 & 1.0001 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.0001 \\ 2.0001 \end{bmatrix}$$

avec solutions exacte et approchée

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \hat{\mathbf{x}} = \begin{bmatrix} 2.0000 \\ 0.0001 \end{bmatrix}.$$

RÉSOLUTION. Le résiduel $\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}$ de la solution approchée (à 4 décimales) est très petit :

$$\begin{aligned} \mathbf{r} &= \begin{bmatrix} 2.0001 \\ 2.0001 \end{bmatrix} - \begin{bmatrix} 1.0001 & 1 \\ 1 & 1.0001 \end{bmatrix} \begin{bmatrix} 2.0000 \\ 0.0001 \end{bmatrix} \\ &= \begin{bmatrix} 2.0001 \\ 2.0001 \end{bmatrix} - \begin{bmatrix} 2.0003 \\ 2.0001 \end{bmatrix} = \begin{bmatrix} -0.0002 \\ 0.0000 \end{bmatrix}. \end{aligned}$$

Cependant, l'erreur relative en $\hat{\mathbf{x}}$,

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \frac{(1.0000 + 0.9999)}{1 + 1} \approx 1,$$

est 100% parce que le système est très mal conditionné :

$$A^{-1} = \frac{1}{0.0002} \begin{bmatrix} 1.0001 & -1.0000 \\ -1.0000 & 1.0001 \end{bmatrix} = \begin{bmatrix} 5000.5 & -5000.0 \\ -5000.0 & 5000.5 \end{bmatrix},$$

et

$$\kappa_1(A) = (1.0001 + 1.0000)(5000.5 + 5000.0) = 20\,002. \quad \square$$

On obtient la norme l_1 de A de l'exemple précédent et le conditionnement de A dans la norme l_1 par les commandes suivantes de Matlab numérique :

```
>> A = [1.0001 1; 1 1.0001];
>> N1 = norm(A,1)
N1 = 2.0001
>> K1 = cond(A,1)
K1 = 2.0001e+04
```

11.4. Méthodes de récurrence

On peut résoudre les systèmes linéaires par récurrence surtout les très grands systèmes. La **récurrence de Gauss–Seidel** utilise la valeur la plus récente de chaque variable. On explique la méthode par un exemple.

EXEMPLE 11.6. Itérer deux fois la méthode de Gauss–Seidel sur le système

$$\begin{aligned} 4x_1 + 2x_2 + x_3 &= 14, & x_1^{(0)} &= 1, \\ x_1 + 5x_2 - x_3 &= 10, & x_2^{(0)} &= 1, \\ x_1 + x_2 + 8x_3 &= 20, & x_3^{(0)} &= 1. \end{aligned}$$

RÉSOLUTION. La récurrence de Gauss-Seidel va converger puisque la matrice du système est à diagonale dominante. On réécrit le système sous forme de récurrence:

$$\begin{aligned} x_1^{(n+1)} &= \frac{1}{4}(14 - 2x_2^{(n)} - x_3^{(n)}), & x_1^{(0)} &= 1, \\ x_2^{(n+1)} &= \frac{1}{5}(10 - x_1^{(n+1)} + x_3^{(n)}), & x_2^{(0)} &= 1, \\ x_3^{(n+1)} &= \frac{1}{8}(20 - x_1^{(n+1)} - x_2^{(n+1)}), & x_3^{(0)} &= 1. \end{aligned}$$

Avec $n = 0$, on a

$$\begin{aligned} x_1^{(1)} &= \frac{1}{4}(14 - 2 - 1) = \frac{11}{4} = 2.75 \\ x_2^{(1)} &= \frac{1}{5}(10 - 2.75 + 1) = 1.65 \\ x_3^{(1)} &= \frac{1}{8}(20 - 2.75 - 1.65) = 1.95. \end{aligned}$$

Avec $n = 1$:

$$\begin{aligned} x_1^{(2)} &= \frac{1}{4}(14 - 2 \times 1.65 - 1.95) = 2.1875 \\ x_2^{(2)} &= \frac{1}{5}(10 - 2.1875 + 1.95) = 1.9525 \\ x_3^{(2)} &= \frac{1}{8}(20 - 2.1875 - 1.9525) = 1.9825 \quad \square \end{aligned}$$

La récurrence de Gauss-Seidel pour résoudre le système $A\mathbf{x} = \mathbf{b}$ est de la forme :

$$\mathbf{x}^{(m+1)} = D^{-1}(\mathbf{b} - L\mathbf{x}^{(m+1)} - U\mathbf{x}^{(m)}), \quad \text{avec démarrage approprié } \mathbf{x}^{(0)},$$

où la matrice A est la somme de trois matrices,

$$A = D + L + U,$$

avec D diagonale, L triangulaire inférieure stricte, et U triangulaire supérieure stricte.

On a programmé cet algorithme en Matlab pour faire $k = 5$ itérations sur le système :

```
A = [7 1 -1;1 11 1;-1 1 9]; b = [3 0 -17]';
D = diag(A); L = tril(A,-1); U = triu(A,1);
m = size(b,1); % nombre de ligne de b
x = ones(m,1); % valeur de demarrage
y = zeros(m,1); % mise en memoire temporaire
k = 5; % nombre d'iterations
for j = 1:k
uy = U*x(:,j);
for i = 1:m
y(i) = (1/D(i))*(b(i)-L(i,:)*y-uy(i));
end
x = [x,y];
end
x
x =
1.0000    0.4286    0.1861    0.1380    0.1357    0.1356
```

$$\begin{array}{cccccc} 1.0000 & -0.1299 & 0.1492 & 0.1588 & 0.1596 & 0.1596 \\ 1.0000 & -1.8268 & -1.8848 & -1.8912 & -1.8915 & -1.8916 \end{array}$$

Il est important de réarranger la matrice des coefficients d'un système linéaire le plus possible à diagonale dominante pour assurer et améliorer la convergence de la récurrence de Gauss-Seidel.

EXEMPLE 11.7. Réarranger le système

$$\begin{array}{rcl} 2x_1 & + & 10x_2 & - & x_3 & = & -32 \\ -x_1 & + & 2x_2 & - & 15x_3 & = & 17 \\ 10x_1 & - & x_2 & + & 2x_3 & = & 35 \end{array}$$

pour assurer la convergence de la récurrence de Gauss-Seidel.

RÉSOLUTION. On obtient un système à diagonale dominante si la dernière équation devient la 1ère,

$$\begin{array}{rcl} 10x_1 & - & x_2 & + & 2x_3 & = & 35 \\ 2x_1 & + & 10x_2 & - & x_3 & = & -32 \\ -x_1 & + & 2x_2 & - & 15x_3 & = & 17 \end{array} \quad \square$$

La **récurrence de Jacobi** résout le système $A\mathbf{x} = \mathbf{b}$ by the following simultaneous iterative scheme:

$$\mathbf{x}^{(m+1)} = D^{-1} \left(\mathbf{b} - L\mathbf{x}^{(m)} - U\mathbf{x}^{(m)} \right), \quad \text{avec démarrage approprié } \mathbf{x}^{(0)},$$

où les matrices D , L et U déjà définies.

On formule la récurrence de Jacobi pour l'exemple 11.6 :

$$\begin{array}{rcl} x_1^{(n+1)} & = & \frac{1}{4}(14 & - & 2x_2^{(n)} & - & x_3^{(n)}), & x_1^{(0)} & = & 1, \\ x_2^{(n+1)} & = & \frac{1}{5}(10 & - & x_1^{(n)} & & + & x_3^{(n)}), & x_2^{(0)} & = & 1, \\ x_3^{(n+1)} & = & \frac{1}{8}(20 & - & x_1^{(n)} & - & x_2^{(n)} & &), & x_3^{(0)} & = & 1. \end{array}$$

On énonce sans démonstration trois théorèmes sur la convergence des récurrences de Jacobi et de Gauss-Seidel.

THÉORÈME 11.3. *Si la matrice A est à diagonale dominante, alors les récurrences de Jacobi et de Gauss-Seidel convergent.*

THÉORÈME 11.4. *Soit une matrice $A \in \mathbb{R}^{n \times n}$ telle que $a_{ii} > 0$ et $a_{ij} \leq 0$ pour $i \neq j$, $i, j = 1, 2, \dots, n$. Si la récurrence de Jacobi converge, alors la récurrence de Gauss-Seidel converge plus rapidement. Si la récurrence de Jacobi diverge, alors la récurrence de Gauss-Seidel diverge plus rapidement.*

THÉORÈME 11.5. *Si $A \in \mathbb{R}^{n \times n}$ est définie positive, alors la récurrence de Gauss-Seidel converge pour tout $\mathbf{x}^{(0)}$.*

11.5. Systèmes surdéterminés

Un système linéaire est **surdéterminé** si le nombre d'équations dépasse le nombre d'inconnues.

On considère l'approximation d'une courbe au sens des moindres carrés. Soit N points,

$$(x_1, y_1), \quad (x_2, y_2), \quad \dots, \quad (x_N, y_N).$$

On cherche une fonction $f(x)$ qui approche y_i en x_i :

$$f(x_i) \approx y_i, \quad i = 1, 2, \dots, N.$$

On pose

$$f(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots + a_n\varphi_n(x),$$

pour un choix approprié de n fonctions, $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$, et l'on minimise la forme quadratique :

$$Q(a_0, a_1, \dots, a_n) = \sum_{i=1}^N (f(x_i) - y_i)^2.$$

Souvent, $N \gg n + 1$. La forme quadratique est non-dégénérée si les fonctions $\varphi_j(x)$ sont "linéairement indépendantes". Dans ce cas, les 1ères dérivées de Q

$$\frac{\partial Q}{\partial a_j} = 0, \quad j = 0, 1, 2, \dots, n,$$

s'annulent au minimum de Q , On écrit Q au long,

$$Q = \sum_{i=1}^N (a_0\varphi_0(x_i) + \cdots + a_n\varphi_n(x_i) - y_i)^2,$$

et l'on annule les 1ères dérivées de Q par rapport à a_j :

$$\frac{\partial Q}{\partial a_j} = 2 \sum_{i=1}^N (a_0\varphi_0(x_i) + \cdots + a_n\varphi_n(x_i) - y_i)\varphi_j(x_i) = 0.$$

On a donc un système linéaire symétrique d'ordre $(n + 1)$:

$$\begin{aligned} \begin{bmatrix} \sum \varphi_0(x_i)\varphi_0(x_i) & \sum \varphi_1(x_i)\varphi_0(x_i) & \cdots & \sum \varphi_n(x_i)\varphi_0(x_i) \\ \vdots & & & \vdots \\ \sum \varphi_0(x_i)\varphi_n(x_i) & \sum \varphi_1(x_i)\varphi_n(x_i) & \cdots & \sum \varphi_n(x_i)\varphi_n(x_i) \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix} \\ = \begin{bmatrix} \sum \varphi_0(x_i)y_i \\ \vdots \\ \sum \varphi_n(x_i)y_i \end{bmatrix}, \quad (11.3) \end{aligned}$$

où l'on somme sur les i de 1 à N . Écrivant la matrice A d'ordre $N \times (n + 1)$, et le vecteur \mathbf{y} d'ordre N sous la forme

$$A = \begin{bmatrix} \varphi_0(x_1) & \varphi_1(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & & & \vdots \\ \varphi_0(x_N) & \varphi_1(x_N) & \cdots & \varphi_n(x_N) \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix},$$

on voit que le système (11.3) d'ordre $n + 1$ s'écrit sous la forme

$$A^T A \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix} = A^T \mathbf{y}.$$

Ces équations s'appelle *équations normales*.

Dans le cas de la régression linéaire, on a

$$\varphi_0(x) = 1, \quad \varphi_1(x) = x,$$

et les équations normales sont

$$\begin{bmatrix} N & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i y_i \end{bmatrix}.$$

C'est l'approximation des données par une droite au sens des moindres carrés.

Dans le cas de la régression quadratique, on a la base

$$\varphi_0(x) = 1, \quad \varphi_1(x) = x, \quad \varphi_2(x) = x^2,$$

et les équations normales

$$\begin{bmatrix} N & \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^3 \\ \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i y_i \\ \sum_{i=1}^N x_i^2 y_i \end{bmatrix}.$$

C'est l'approximation des données par une parabole au sens des moindres carrés.

EXEMPLE 11.8. Approximer les données :

| | | | | | |
|-------|---|---|---|---|---|
| i | 1 | 2 | 3 | 4 | 5 |
| x_i | 0 | 1 | 2 | 4 | 6 |
| y_i | 3 | 1 | 0 | 1 | 4 |

au sens des moindres carrés par la parabole

$$f(x) = a_0 + a_1 x + a_2 x^2.$$

RÉSOLUTION. (a) **Lla solution analytique.**— On a les équations normales :

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 4 & 6 \\ 0 & 1 & 4 & 16 & 36 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 6 & 36 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 4 & 6 \\ 0 & 1 & 4 & 16 & 36 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix},$$

c'est-à-dire

$$\begin{bmatrix} 5 & 13 & 57 \\ 13 & 57 & 289 \\ 57 & 289 & 1569 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 9 \\ 29 \\ 161 \end{bmatrix},$$

ou

$$N\mathbf{a} = \mathbf{b}.$$

Par la décomposition de Cholesky $N = GG^T$ on a

$$G = \begin{bmatrix} 2.2361 & 0 & 0 \\ 5.8138 & 4.8166 & 0 \\ 25.4921 & 29.2320 & 8.0430 \end{bmatrix}.$$

On obtient a solution \mathbf{a} par substitutions prograde et rétrograde avec $G\mathbf{w} = \mathbf{b}$ et $G^T\mathbf{a} = \mathbf{w}$,

$$\begin{aligned} a_0 &= 2.8252 \\ a_1 &= -2.0490 \\ a_2 &= 0.3774. \end{aligned}$$

(b) **La solution par Matlab numérique.**—

```

x = [0 1 2 4 6]';
A = [x.^0 x x.^2];
y = [3 1 0 1 4]';
a = (A'*A \ (A'*y))'
a = 2.8252 -2.0490 0.3774

```

Le résultat est tracé dans la Fig. 11.1 □

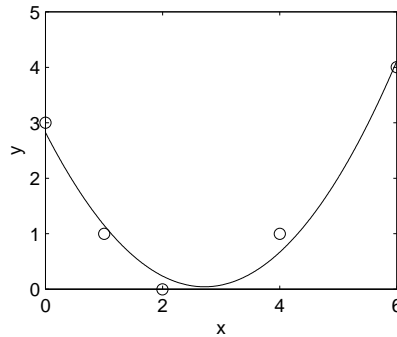


FIGURE 11.1. Approximation quadratique au sens des moindres carrés de l'exemple 11.8.

11.6. Valeurs propres et vecteurs propres

Une *valeur propre*, ou *valeur caractéristique*, d'une matrice $A \in \mathbb{R}^{n \times n}$, ou $\in \mathbb{C}^{n \times n}$, est un nombre réel ou complexe telle que l'équation vectorielle

$$A\mathbf{x} = \lambda\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n \text{ ou } \mathbb{C}^n, \quad (11.4)$$

admet un solution non nulle, $\mathbf{x} \neq 0$, appelée *vecteur propre*. On récrit (11.4) sous la forme

$$(A - \lambda I)\mathbf{x} = 0, \quad (11.5)$$

où I est la matrice identité d'ordre n . Cette équation admet une solution non nulle \mathbf{x} si et seulement si le *déterminant caractéristique*

$$\det(A - \lambda I) = 0, \quad (11.6)$$

est nul, c'est-à-dire λ est un zéro du polynôme caractéristique de A .

11.6.1. Disques de Gershgorin. Le théorème d'inclusion de Gershgorin affirme que chaque valeur propre de A est dans un disque de Gershgorin.

THÉORÈME 11.6 (Théorème de Gershgorin). *Soit λ une valeur propre d'une matrice $A = (a_{ij})$ quelconque d'ordre n . Alors pour un certain i , $1 \leq i \leq n$, on a l'inégalité :*

$$|a_{ii} - \lambda| \leq |a_{i1}| + |a_{i2}| + \cdots + |a_{i,i-1}| + |a_{i,i+1}| + \cdots + |a_{in}|. \quad (11.7)$$

PROOF. Soit \mathbf{x} un vecteur propre associé à la valeur propre λ :

$$(A - \lambda I)\mathbf{x} = 0, \quad (11.8)$$

et x_i la composante maximum de \mathbf{x} en module. Alors $|x_j/x_i| \leq 1$ pour $j = 1, 2, \dots, n$. L'équation vectorielle (11.8) est un système de n équations et la i ème équation est

$$a_{i1}x_1 + \dots + a_{i,i-1}x_{i-1} + (a_{ii} - \lambda)x_i + a_{i,i+1}x_{i+1} + \dots + a_{in}x_n = 0.$$

On divise cette équation par x_i et on l'on résout pour $a_{ii} - \lambda$:

$$a_{ii} - \lambda = -a_{i1}\frac{x_1}{x_i} - \dots - a_{i,i-1}\frac{x_{i-1}}{x_i} - a_{i,i+1}\frac{x_{i+1}}{x_i} - \dots - a_{in}\frac{x_n}{x_i}.$$

On prend le module des deux membres et l'on applique l'inégalité du triangle $|a + b| \leq |a| + |b|$ (où a et b sont de nombres complexes quelconques). Alors, par le choix de i , on a

$$\left| \frac{x_1}{x_i} \right| \leq 1, \quad \dots, \quad \left| \frac{x_n}{x_i} \right| \leq 1,$$

et (11.7) suit. □

EXEMPLE 11.9. Au moyen du théorème de Gershgorin, déterminer et tracer les disques de Gershgorin D_k qui contiennent les valeurs propres de la matrice

$$A = \begin{bmatrix} -3 & 0.5i & -i \\ 1-i & 1+i & 0 \\ 0.1i & 1 & -i \end{bmatrix}.$$

RÉSOLUTION. Les centres c_i et les rayons r_i des disques sont

$$\begin{aligned} c_1 &= -3, & r_1 &= |0.5i| + |-i| = 1.5 \\ c_2 &= 1+i, & r_2 &= |1-i| + |0| = \sqrt{2} \\ c_3 &= -i, & r_3 &= |0.1i| + 1 = 1.1 \end{aligned}$$

tel qu'indiqué dans la Fig. 11.2. □

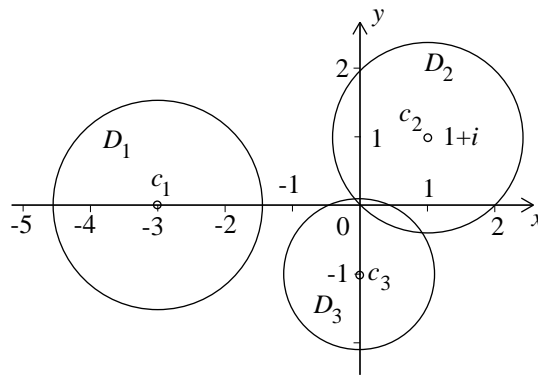


FIGURE 11.2. Disques de Gershgorin pour l'exemple 11.9.

Les valeurs propres de la matrice A de l'exemple 11.9, trouvées par ordinateur (V. l'exemple 11.10) sont

$$-3.2375 - 0.1548i, \quad 1.0347 + 1.1630i, \quad 0.2027 - 1.0082i.$$

11.6.2. La méthode de la puissance. On peut employer la *méthode de la puissance* pour trouver la valeur propre de A la plus grande en module et le vecteur propre associé. On dérive cette méthode.

Pour simplifier la dérivation on suppose que A admet n vecteurs propres $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$ linéairement indépendants aux valeurs propres associées $\lambda_1, \lambda_2, \dots, \lambda_n$, décroissantes en module :

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|.$$

On peut alors représenter un vecteur \mathbf{x} quelconque sous la forme

$$\mathbf{x} = a_1 \mathbf{z}_1 + a_2 \mathbf{z}_2 + \dots + a_n \mathbf{z}_n.$$

On applique A^k sur \mathbf{x} :

$$\begin{aligned} A^k \mathbf{x} &= a_1 \lambda_1^k \mathbf{z}_1 + a_2 \lambda_2^k \mathbf{z}_2 + \dots + a_n \lambda_n^k \mathbf{z}_n \\ &= \lambda_1^k \left[a_1 \mathbf{z}_1 + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{z}_2 + \dots + a_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \mathbf{z}_n \right] \\ &\rightarrow \lambda_1^k a_1 \mathbf{z}_1 = \mathbf{y} \quad \text{quand } k \rightarrow \infty. \end{aligned}$$

Donc $A\mathbf{y} = \lambda_1 \mathbf{y}$. En pratique, on étalonne les vecteurs pour éviter les dépassements de capacité :

$$\begin{aligned} A\mathbf{x}^{(0)} &= \mathbf{x}^{(1)}, & \mathbf{u}^{(1)} &= \frac{\mathbf{x}^{(1)}}{\|\mathbf{x}^{(1)}\|_\infty}, \\ A\mathbf{u}^{(1)} &= \mathbf{x}^{(2)}, & \mathbf{u}^{(2)} &= \frac{\mathbf{x}^{(2)}}{\|\mathbf{x}^{(2)}\|_\infty}, \\ & \vdots & & \\ A\mathbf{u}^{(n)} &= \mathbf{x}^{(n+1)} \\ & \approx \lambda_1 \mathbf{u}^{(n)}. \end{aligned}$$

EXEMPLE 11.10. Trouver la valeur propre de module maximum de la matrice

$$\begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix}$$

et le vecteur propre associé.

RÉSOLUTION. Posons $\mathbf{x}^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Alors

$$\begin{aligned} \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= \begin{bmatrix} 5 \\ 7 \end{bmatrix} = \mathbf{x}^{(1)}, & \mathbf{u}^{(1)} &= \begin{bmatrix} 5/7 \\ 1 \end{bmatrix}, \\ \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 5/7 \\ 1 \end{bmatrix} &= \begin{bmatrix} 4.14 \\ 6.43 \end{bmatrix} = \mathbf{x}^{(2)}, & \mathbf{u}^{(2)} &= \begin{bmatrix} 0.644 \\ 1 \end{bmatrix}, \\ \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 0.644 \\ 1 \end{bmatrix} &= \begin{bmatrix} 3.933 \\ 6.288 \end{bmatrix} = \mathbf{x}^{(3)}, & \mathbf{u}^{(3)} &= \begin{bmatrix} 0.6254 \\ 1 \end{bmatrix}. \end{aligned}$$

Donc

$$\lambda_1 \approx 6.288, \quad \mathbf{x}_1 \approx \begin{bmatrix} 0.6254 \\ 1 \end{bmatrix}. \quad \square$$

La commande `eig` de Matlab numérique calcule les valeurs propres et les vecteurs propres d'une matrice numérique. Par exemple,

```
>> A = [3 2; 2 5];
>> [X,D] = eig(A)
X =
    0.8507    0.5257
   -0.5257    0.8507
D =
    1.7639         0
         0    6.2361
```

Les vecteurs propres de A sont les colonnes de la matrice X et les vecteurs propres sont les éléments de la matrice diagonale D . La commande numérique `eig` emploie l'algorithme QR avec déplacements expliqué à la section 11.8.

11.6.3. La méthode de la puissance inverse. La *méthode de la puissance inverse* peut calculer n'importe quelle valeur propre d'une matrice quelconque $A \in \mathbb{R}^{n \times n}$ ou $\in \mathbb{C}^{n \times n}$. On la dérive ici sous l'hypothèse simplificatrice que A admet n vecteurs propres $\mathbf{z}_1, \dots, \mathbf{z}_n$ indépendants et que λ est près de λ_1 .

On a :

$$(A - \lambda I)\mathbf{x}^{(1)} = \mathbf{x}^{(0)} = a_1 \mathbf{z}_1 + \dots + a_n \mathbf{z}_n,$$

$$\mathbf{x}^{(1)} = a_1 \frac{1}{\lambda_1 - \lambda} \mathbf{z}_1 + a_2 \frac{1}{\lambda_2 - \lambda} \mathbf{z}_2 + \dots + a_n \frac{1}{\lambda_n - \lambda} \mathbf{z}_n,$$

et par récurrence,

$$\begin{aligned} \mathbf{x}^{(k)} &= a_1 \frac{1}{(\lambda_1 - \lambda)^k} \left[\mathbf{z}_1 + a_2 \left(\frac{\lambda_1 - \lambda}{\lambda_2 - \lambda} \right)^k \mathbf{z}_2 + \dots + a_n \left(\frac{\lambda_1 - \lambda}{\lambda_n - \lambda} \right)^k \mathbf{z}_n \right] \\ &\rightarrow a_1 \frac{1}{(\lambda_1 - \lambda)^k} \mathbf{z}_1, \quad \text{avec } k \rightarrow \infty, \end{aligned}$$

puisque

$$\left| \frac{\lambda_1 - \lambda}{\lambda_j - \lambda} \right| < 1, \quad j \neq 1.$$

Donc, la suite $\mathbf{x}^{(k)}$ converge dans la direction de \mathbf{z}_1 . En pratique, on normalise les vecteurs $\mathbf{x}^{(k)}$ et l'on résout le système

$$(A - \lambda I)\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$$

par décomposition LU selon l'algorithme suivant :

Choisir $\mathbf{x}^{(0)}$

For $k = 1, 2, 3, \dots$, do

Résoudre

$(A - \lambda I)\mathbf{y}^{(k)} = \mathbf{x}^{(k-1)}$ par décomposition LU avec pivotage partiel.

$\mathbf{x}^{(k)} = \mathbf{y}^{(k)} / \|\mathbf{y}^{(k)}\|_\infty$

Stop if $\|(A - \lambda I)\mathbf{x}^{(k)}\|_\infty < c\epsilon\|A\|_\infty$, où c est une constante de l'ordre de un et ϵ est l'epsilon machine.

11.7. La décomposition QR

La décomposition QR est une puissante méthode pour résoudre des systèmes mal conditionnés ou surdéterminés,

$$A\mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{m \times n}, \quad m \geq n.$$

Cette méthode décompose la matrice A sous la forme

$$A = QR,$$

où Q est orthogonale ou unitaire et R est triangulaire supérieure. Alors,

$$\|A\mathbf{x} - \mathbf{b}\|_2 = \|QR\mathbf{x} - \mathbf{b}\|_2 = \|R\mathbf{x} - Q^T\mathbf{b}\|_2.$$

Si A est de rang plein, c'est-à-dire le rang de A est égal à n , on peut écrire :

$$R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}, \quad Q^T\mathbf{b} = \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix},$$

où $R_1 \in \mathbb{R}^{n \times n}$, $0 \in \mathbb{R}^{(m-n) \times n}$, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{d} \in \mathbb{R}^{m-n}$ et R_1 régulière et triangulaire supérieure.

Alors, on obtient la solution au sens des moindres carrés,

$$\mathbf{x} = R_1^{-1}\mathbf{c},$$

par substitution rétrograde sur

$$R_1\mathbf{x} = \mathbf{c}.$$

Le *résidu* est

$$\rho = \min_{\mathbf{x} \in \mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|_2 = \|\mathbf{d}\|_2.$$

La décomposition QR transforme une matrice A en une matrice triangulaire supérieure au moyen de $n-1$ réflexions de Householder. La k ième réflexion annule les éléments sous la diagonale de la k ième colonne. Par exemple, pour annuler les composantes x_2, x_3, \dots, x_n du vecteur $\mathbf{x} \in \mathbb{R}^n$, on applique la réflexion de Householder

$$P = I - 2 \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}},$$

avec

$$\mathbf{v} = \mathbf{x} + \text{sign}(x_1)\|\mathbf{x}\|_2\mathbf{e}_1, \quad \text{où } \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Dans ce cas,

$$P \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} -\|\mathbf{x}\|_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

La matrice P est symétrique et orthogonale et sa propre inverse :

$$P^T = P = P^{-1}.$$

Pour minimiser le nombre d'opérations en virgule flottante et l'allocation de la mémoire, on calcule d'abord le scalaire

$$s = 2/\mathbf{v}^T\mathbf{v}$$

puis le produit

$$P\mathbf{x} = \mathbf{x} - s(\mathbf{v}^T \mathbf{x})\mathbf{v}$$

en tenant compte de la structure spéciale de la matrice P . On ne garde en mémoire que le nombre s et le vecteur \mathbf{v} .

Les logiciels emploient systématiquement la décomposition QR pour résoudre les systèmes surdéterminés. Il en est ainsi de la division à gauche de Matlab \ sur les systèmes surdéterminés. La commande numérique `qr` de Matlab fait la décomposition QR d'une matrice :

```
>> A = [1 2 3; 4 5 6; 7 8 9];
>> [Q,R] = qr(A)
Q =
  -0.1231    0.9045    0.4082
  -0.4924    0.3015   -0.8165
  -0.8616   -0.3015    0.4082
R =
  -8.1240   -9.6011  -11.0782
         0    0.9045    1.8091
         0         0   -0.0000
```

On voit que A est singulière puisque $r_{33} = 0$.

11.8. L'algorithme QR

L'algorithme QR emploie une suite de décompositions QR :

$$\begin{aligned} A &= Q_1 R_1 \\ A_1 &= R_1 Q_1 = Q_2 R_2 \\ A_2 &= R_2 Q_2 = Q_3 R_3 \\ &\vdots \end{aligned}$$

pour calculer les valeurs propres de A , puisque A_n converge sur une matrice triangulaire supérieure ou quasi-supérieure avec les valeurs propres réelles sur la diagonale et les valeurs propres sur des blocs diagonaux d'ordre 2. On accélère grandement la convergence par des déplacements simples ou doubles ou d'autres déplacements.

On a rarement besoin de toutes les valeurs propres d'une grande matrice d'ordre $n \geq 100$. On emploie plutôt la méthode de Lanczos pour calculer les valeurs propres désirées.

On popularise la méthode de Jacobi pour valeurs propres de matrices symétriques parce qu'elle est parallélisable sur ordinateurs parallèles.

11.9. La décomposition selon les valeurs singulières

La *décomposition d'une matrice selon les valeurs singulières* est un outil très puissant en calcul matriciel mais plus coûteux en temps que les méthodes antérieures. Une matrice quelconque $A \in \mathbb{R}^{m \times n}$, disons, avec $m \geq n$, admet la factorisation de la forme

$$A = U\Sigma V^T,$$

où les matrices $U \in \mathbb{R}^{m \times m}$ et $V \in \mathbb{R}^{n \times n}$ sont orthogonales. La matrice $\Sigma \in \mathbb{R}^{m \times n}$ est diagonale et ses éléments σ_i , en ordre décroissant,

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0,$$

sont les valeurs singulières de A . Si $A \in \mathbb{R}^{n \times n}$. On voit que

$$\|A\|_2 = \sigma_1, \quad \|A^{-1}\|_2 = 1/\sigma_n.$$

Si $A \in \mathbb{C}^{m \times n}$ est complexe, alors U et V sont unitaires et l'on remplace la transposée V^T par la transposée hermitienne

$$V^H = \bar{V}^T.$$

Le rang de A est égal au nombre de valeurs singulières non-nulles.

La commande `svd` de Matlab numérique fait la décomposition selon les valeurs singulières d'une matrice A :

`A = [1 2 3; 4 5 6; 7 8 9];`

`[U,S,V] = svd(A)`

`U =`

```

    0.2148    0.8872   -0.4082
    0.5206    0.2496    0.8165
    0.8263   -0.3879   -0.4082
```

`S =`

```

   16.8481         0         0
         0    1.0684         0
         0         0    0.0000
```

`V =`

```

    0.4797   -0.7767    0.4082
    0.5724   -0.0757   -0.8165
    0.6651    0.6253    0.4082
```

Les éléments de la matrice diagonale S sont les valeurs singulières de A . La norme l_2 de A est égale à $\|A\|_2 = \sigma_1 = 16.8481$. Puisque $\sigma_3 = 0$, A est singlière.

Si A est symétrique, $A^T = A$, hermitienne $A^H = A$ ou même *normale*, $AA^H = A^H A$, alors les valeurs propres de A en module sont les valeurs singulières de A .

THÉORÈME 11.7 (Décomposition de Schur). *Une matrice carrée A quelconque admet la décomposition de Schur*

$$A = UTU^H,$$

où les éléments de la diagonale de la matrice triangulaire supérieure T sont les valeurs propres de A et la matrice U est unitaire.

Si A est normale, alors T est diagonale.

THÉORÈME 11.8. *Une matrice A est normale si et seulement elle admet la décomposition de Schur*

$$A = UDU^H,$$

où la matrice diagonale D contient les valeurs propres de A et les colonnes de la matrice unitaire U sont les vecteurs propres de A .

Partie 3

Exercices suivis des solutions des
exercices marqués d'un astérisque

Exercices pour équadifs et transf. de Laplace

Les solutions des exercices marqués d'un astérisque se trouvent au chapitre 14.

Exercices pour le chapitre premier

Résoudre les équations différentielles séparables.

1.1. $y' = 2xy^2$.

1.2. $y' = \frac{xy}{x^2 - 1}$.

***1.3.** $(1 + x^2)y' = \cos^2 y$.

1.4. $(1 + e^x)yy' = e^x$.

1.5. $y' \sin x = y \ln y$.

1.6. $(1 + y^2) dx + (1 + x^2) dy = 0$.

5.56.

Résoudre les problèmes à valeur initiale et tracer les solutions

1.7. $y' \sin x - y \cos x = 0, \quad y(\pi/2) = 1$.

1.8. $x \sin y dx + (x^2 + 1) \cos y dy = 0, \quad y(1) = \pi/2$.

Résoudre les équations différentielles.

1.9. $(x^2 - 3y^2) dx + 2xy dy = 0$.

1.10. $(x + y) dx - x dy = 0$.

***1.11.** $xy' = y + \sqrt{y^2 - x^2}$.

1.12. $xy' = y + x \cos^2(y/x)$.

Résoudre les problèmes à valeur initiale.

1.13. $(2x - 5y) dx + (4x - y) dy = 0, \quad y(1) = 4$.

1.14. $(3x^2 + 9xy + 5y^2) dx - (6x^2 + 4xy) dy = 0, \quad y(2) = -6$.

1.15. $yy' = -(x + 2y), \quad y(1) = 1$.

1.16. $(x^2 + y^2) dx - 2xy dy = 0, \quad y(1) = 2$.

Résoudre les équations différentielles.

1.17. $x(2x^2 + y^2) + y(x^2 + 2y^2)y' = 0$.

1.18. $(3x^2y^2 - 4xy)y' + 2xy^3 - 2y^2 = 0$.

$$1.19. (\sin xy + xy \cos xy) dx + x^2 \cos xy dy = 0.$$

$$1.20. \left(\frac{\sin 2x}{y} + x \right) dx + \left(y - \frac{\sin^2 x}{y^2} \right) dy = 0.$$

Résoudre les problèmes à valeur initiale.

$$*1.21. (2xy - 3) dx + (x^2 + 4y) dy = 0, \quad y(1) = 2.$$

$$1.22. \frac{2x}{y^3} dx + \frac{y^2 - 3x^2}{y^4} dy = 0, \quad y(1) = 1.$$

$$1.23. (y e^x + 2 e^x + y^2) dx + (e^x + 2xy) dy = 0, \quad y(0) = 6.$$

$$1.24. (2x \cos y + 3x^2 y) dx + (x^3 - x^2 \sin y - y) dy = 0, \quad y(0) = 2.$$

Résoudre les équations différentielles.

$$*1.25. (x + y^2) dx - 2xy dy = 0.$$

$$1.26. (x^2 - 2y) dx + x dy = 0.$$

$$1.27. (x^2 - y^2 + x) dx + 2xy dy = 0.$$

$$1.28. (1 - x^2 y) dx + x^2 (y - x) dy = 0.$$

$$1.29. (1 - xy)y' + y^2 + 3xy^3 = 0.$$

$$1.30. (2xy^2 - 3y^3) dx + (7 - 3xy^2) dy = 0.$$

$$1.31. (2x^2 y - 2y + 5) dx + (2x^3 + 2x) dy = 0.$$

$$1.32. (x + \sin x + \sin y) dx + \cos y dy = 0.$$

$$1.33. y' + \frac{2}{x} y = 12.$$

$$1.34. y' + \frac{2x}{x^2 + 1} y = x.$$

$$1.35. x(\ln x)y' + y = 2 \ln x.$$

$$1.36. xy' + 6y = 3x + 1.$$

Résoudre les problèmes à valeur initiale.

$$1.37. y' + 3x^2 y = x^2, \quad y(0) = 2.$$

$$1.38. xy' - 2y = 2x^4, \quad y(2) = 8.$$

$$*1.39. y' + y \cos x = \cos x, \quad y(0) = 1.$$

$$1.40. y' - y \tan x = \frac{1}{\cos^3 x}, \quad y(0) = 0.$$

Trouver les trajectoires orthogonale des familles de courbes. Dans chaque cas, tracer quelques courbes des deux familles sur le même repère.

$$1.41. x^2 + y^2/4 = c.$$

$$1.42. y = e^x + c.$$

$$1.43. y^2 + 2x = c.$$

$$1.44. y = \arctan x + c.$$

1.45. $x^2 - y^2 = c^2$.

1.46. $y^2 = cx^3$.

1.47. $e^x \cos y = c$.

1.48. $y = \ln x + c$.

Tracer le champs des tangentes et tracer quelques courbes solutions approchées.

1.49. $y' = 2y/x$.

1.50. $y' = -x/y$.

1.50. $y' = -xy$.

1.51. $9yy' + x = 0$.

Exercices pour le chapitre 2

Résoudre les équations différentielles.

2.1. $y'' - 3y' + 2y = 0$.

2.2. $y'' + 2y' + y = 0$.

*2.3. $y'' - 9y' + 20y = 0$.

Résoudre les problèmes à valeur initiale $y(x_0) = y_0$, et tracer les solutions $y(x)$ pour $x \geq x_0$.

2.4. $y'' + y' + \frac{1}{4}y = 0$, $y(2) = 1$, $y'(2) = 1$.

2.5. $y'' + 9y = 0$, $y(0) = 0$, $y'(0) = 1$.

2.6. $y'' - 4y' + 3y = 0$, $y(0) = 6$, $y'(0) = 0$.

2.7. $y'' - 2y' + 3y = 0$, $y(0) = 1$, $y'(0) = 3$.

2.8. $y'' + 2y' + 2y = 0$, $y(0) = 2$, $y'(0) = -3$.

Trouver l'amplitude et la période de l'oscillateur non amorti.

2.9. $y'' + 4y = 0$, $y(0) = 1$, $y'(0) = 2$.

2.10. $y'' + 16y = 0$, $y(0) = 0$, $y'(0) = 1$.

Considérons un oscillateur sous amortissement critique. Trouver une valeur $T \geq 0$ telle que $|y(T)|$ soit maximum, trouver le maximum, et tracer les solutions $y(x)$ pour $x \geq x_0$.

2.11. $y'' + 2y' + y = 0$, $y(0) = 1$, $y'(0) = 1$.

2.12. $y'' + 6y' + 9y = 0$, $y(0) = 0$, $y'(0) = 2$.

Résoudre les équations différentielles d'Euler-Cauchy.

*2.13. $x^2 y'' + 3xy' - 3y = 0$.

2.14. $x^2 y'' - xy' + y = 0$.

2.15. $4x^2y'' + y = 0$.

2.16. $x^2y'' + xy' + 4y = 0$.

Résoudre les problèmes à valeur initiale $y(x_0) = y_0$, et tracer les solutions $y(x)$ pour $x \geq x_0$.

2.17. $x^2y'' + 4xy' + 2y = 0$, $y(1) = 1$, $y'(1) = 2$.

2.18. $x^2y'' + 5xy' + 3y = 0$, $y(1) = 1$, $y'(1) = -5$.

2.19. $x^2y'' - xy' + y = 0$, $y(1) = 1$, $y'(1) = 0$.

2.20. $x^2y'' + \frac{7}{2}xy' - \frac{3}{2}y = 0$, $y(4) = 1$, $y'(4) = 0$.

Exercices pour le chapitre 3

Résoudre les équations différentielles à coefficients constants.

***3.1.** $y''' + 6y'' = 0$.

3.2. $y''' + 3y'' - 4y' - 12y = 0$.

3.3. $y''' - y = 0$.

3.4. $y^{(4)} + y''' - 3y'' - y' + 2y = 0$.

Résoudre les problèmes à valeur initiale et tracer les solutions $y(x)$ pour $x \geq 0$.

3.5. $y''' + 12y'' + 36y' = 0$, $y(0) = 0$, $y'(0) = 1$, $y''(0) = -7$.

3.6. $y^{(4)} - y = 0$, $y(0) = 0$, $y'(0) = 0$, $y''(0) = 0$, $y'''(0) = 1$.

3.7. $y''' - y'' - y' + y = 0$, $y(0) = 0$, $y'(0) = 5$, $y''(0) = 2$.

3.8. $y''' - 2y'' + 4y' - 8y = 0$, $y(0) = 2$, $y'(0) = 0$, $y''(0) = 0$.

Déterminer si les fonctions sont linéairement dépendantes ou indépendantes sur $-\infty < x < +\infty$.

***3.9.** $y_1(x) = x$, $y_2(x) = x^2$, $y_3(x) = 2x - 5x^2$.

3.10. $y_1(x) = 1 + x$, $y_2(x) = x$, $y_3(x) = x^2$.

3.11. $y_1(x) = 2$, $y_2(x) = \sin^2 x$, $y_3(x) = \cos^2 x$.

3.12. $y_1(x) = e^x$, $y_2(x) = e^{-x}$, $y_3(x) = \cosh x$.

Montrer au moyen du wronskien que les fonctions sont linéairement indépendantes sur l'intervalle donné.

***3.13.** e^x , e^{2x} , e^{-x} , $-\infty < x < +\infty$.

3.14. $x + 2$, x^2 , $-\infty < x < +\infty$.

3.15. $x^{1/3}$, $x^{1/4}$, $0 < x < +\infty$.

3.16. x , $x \ln x$, $x^2 \ln x$, $e^{-2} < x < +\infty$.

3.17 Montrer que les fonctions

$$f_1(x) = x^2, \quad f_2(x) = x|x| = \begin{cases} x^2, & x \geq 0, \\ -x^2, & x < 0 \end{cases}$$

sont linéairement indépendantes sur $[-1, 1]$ et calculer leur wronskien. Expliquer vos résultats.

Trouver une seconde solution des équations différentielles si $y_1(x)$ est une solution.

3.18. $xy'' + y' = 0, \quad y_1(x) = \ln x.$

3.19. $x(x-2)y'' - (x^2-2)y' + 2(x-1)y = 0, \quad y_1(x) = e^x.$

3.20. $(1-x^2)y'' - 2xy' = 0, \quad y_1(x) = 1.$

3.21. $(1+2x)y'' + 4xy' - 4y = 0, \quad y_1(x) = e^{-2x}.$

Résoudre les équations différentielles.

3.22. $y'' + 3y' + 2y = 5e^{-2x}.$

3.23. $y'' + y' = 3x^2.$

3.24. $y'' - y' - 2y = 2xe^{-x} + x^2.$

***3.25.** $y'' - y' = e^x \sin x.$

Résoudre les problèmes à valeur initiale et tracer les solutions $y(x)$ pour $x \geq 0$.

3.26. $y'' + y = 2 \cos x, \quad y(0) = 1, \quad y'(0) = 0.$

3.27. $y^{(4)} - y = 8e^x, \quad y(0) = 0, \quad y'(0) = 2, \quad y''(0) = 4, \quad y'''(0) = 6.$

3.28. $y''' + y' = x, \quad y(0) = 0, \quad y'(0) = 1, \quad y''(0) = 0.$

3.29. $y'' + y = 3x^2 - 4 \sin x, \quad y(0) = 0, \quad y'(0) = 1.$

Résoudre les équations différentielles.

3.30. $y'' + y = \frac{1}{\sin x}.$

3.31. $y'' + y = \frac{1}{\cos x}.$

3.32. $y'' + 6y' + 9y = \frac{e^{-3x}}{x^3}.$

3.33. $y'' - 2y' \tan x = 1.$

3.34. $y'' - 2y' + y = \frac{e^x}{x}.$

***3.35.** $y'' + 3y' + 2y = \frac{1}{1+e^x}.$

Résoudre les problèmes à valeur initiale $y(x_0) = y_0$, et tracer les solutions $y(x)$ pour $x \geq x_0$.

3.36. $y'' + y = \tan x, \quad y(0) = 1, \quad y'(0) = 0.$

$$3.37. y'' - 2y' + y = \frac{e^x}{x}, \quad y(1) = e, \quad y'(1) = 0.$$

$$3.38. 2x^2y'' + xy' - 3y = x^{-2}, \quad y(1) = 0, \quad y'(1) = 2.$$

$$3.39. 2x^2y'' + xy' - 3y = 2x^{-3}, \quad y(1) = 0, \quad y'(1) = 3.$$

Exercices pour le chapitre 4

Résoudre les systèmes d'équations différentielles $\mathbf{y}' = A\mathbf{y}$ pour A donné.

$$4.1. A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}.$$

$$4.2. A = \begin{bmatrix} 2 & 0 & 4 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}.$$

$$*4.3. A = \begin{bmatrix} -1 & 1 \\ 4 & -1 \end{bmatrix}.$$

$$4.4. A = \begin{bmatrix} -1 & 1 & 4 \\ -2 & 2 & 4 \\ -1 & 0 & 4 \end{bmatrix}.$$

$$4.5. A = \begin{bmatrix} 1 & 1 \\ -4 & 1 \end{bmatrix}.$$

Résoudre les systèmes d'équations différentielles $\mathbf{y}' = A\mathbf{y} + \mathbf{f}(x)$ pour A et \mathbf{f} donnés.

$$4.6. A = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{f}(x) = \begin{bmatrix} 2e^{-x} \\ -e^{-x} \end{bmatrix}.$$

$$4.7. A = \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}, \quad \mathbf{f}(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

$$4.8. A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}, \quad \mathbf{f}(x) = \begin{bmatrix} 2e^{-x} \\ 3x \end{bmatrix}.$$

$$4.9. A = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}, \quad \mathbf{f}(x) = \begin{bmatrix} e^x \\ -e^x \end{bmatrix}.$$

$$4.10. A = \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \quad \mathbf{f}(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Résoudre les systèmes d'équations différentielles $\mathbf{y}' = A\mathbf{y}$ avec $\mathbf{y}(0) = \mathbf{y}_0$ pour A et \mathbf{y}_0 donnés.

$$4.11. A = \begin{bmatrix} 5 & -1 \\ 3 & 1 \end{bmatrix}, \quad \mathbf{y}_0 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

$$4.12. A = \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix}, \quad \mathbf{y}_0 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

$$4.13. A = \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \quad \mathbf{y}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Exercices pour le chapitre 5

Trouver les transformées de Laplace de $f(t)$.

5.1. $f(t) = -3t + 2$.

5.2. $f(t) = t^2 + at + b$.

5.3. $f(t) = \cos(\omega t + \theta)$.

5.4. $f(t) = \sin(\omega t + \theta)$.

5.5. $f(t) = \cos^2 t$.

5.6. $f(t) = \sin^2 t$.

5.7. $f(t) = 3 \cosh 2t + 4 \sinh 5t$.

5.8. $f(t) = 2 e^{-2t} \sin t$.

5.9. $f(t) = e^{-2t} \cosh t$.

5.10. $f(t) = (1 + 2e^{-t})^2$.

5.11. $f(t) = u(t-1)(t-1)$.

*5.12. $f(t) = u(t-1)t^2$.

5.13. $f(t) = u(t-1) \cosh t$.

5.14. $f(t) = u(t - \pi/2) \sin t$.

Trouver les transformées de Laplace inverses de $F(s)$.

5.15. $F(s) = \frac{4(s+1)}{s^2-16}$.

5.16. $F(s) = \frac{2s}{s^2+3}$.

5.17. $F(s) = \frac{2}{s^2+3}$.

5.18. $F(s) = \frac{4}{s^2-9}$.

5.19. $F(s) = \frac{4s}{s^2-9}$.

5.20. $F(s) = \frac{3s-5}{s^2+4}$.

5.21. $F(s) = \frac{1}{s^2+s-20}$.

5.22. $F(s) = \frac{1}{(s-2)(s^2+4s+3)}$.

*5.23. $F(s) = \frac{2s+1}{s^2+5s+6}$.

5.24. $F(s) = \frac{s^2-5}{s^3+s^2+9s+9}$.

$$5.25. F(s) = \frac{3s^2 + 8s + 3}{(s^2 + 1)(s^2 + 9)}.$$

$$5.26. F(s) = \frac{s - 1}{s^2(s^2 + 1)}.$$

$$5.27. F(s) = \frac{1}{s^4 - 9}.$$

$$5.28. F(s) = \frac{(1 + e^{-2s})^2}{s + 2}.$$

$$5.29. F(s) = \frac{e^{-3s}}{s^2(s - 1)}.$$

$$5.30. F(s) = \frac{\pi}{2} - \arctan \frac{s}{2}.$$

$$5.31. F(s) = \ln \frac{s^2 + 1}{s^2 + 4}.$$

Trouver les transformées de Laplace de $f(t)$.

$$5.32. f(t) = \begin{cases} t, & 0 \leq t < 1, \\ 1, & t \geq 1. \end{cases}$$

$$5.33. f(t) = \begin{cases} 2t + 3, & 0 \leq t < 2, \\ 0, & t \geq 2. \end{cases}$$

$$5.34. f(t) = t \sin 3t.$$

$$5.35. f(t) = t \cos 4t.$$

$$*5.36. f(t) = e^{-t} t \cos t.$$

$$5.37. f(t) = \int_0^t \tau e^{t-\tau} d\tau.$$

$$5.38. f(t) = 1 * e^{-2t}.$$

$$5.39. f(t) = e^{-t} * e^t \cos t.$$

$$5.40. f(t) = \frac{e^t - e^{-t}}{t}.$$

Résoudre les problèmes à valeur initiale par transformation de Laplace et tracer les solutions.

$$5.41. y'' - 6y' + 13y = 0, \quad y(0) = 0, \quad y'(0) = -3.$$

$$5.42. y'' + y = \sin 3t, \quad y(0) = 0, \quad y'(0) = 0.$$

$$5.43. y'' + y = \sin t, \quad y(0) = 0, \quad y'(0) = 0.$$

$$5.44. y'' + y = t, \quad y(0) = 0, \quad y'(0) = 0.$$

$$5.45. y'' + 5y' + 6y = 3e^{-2t}, \quad y(0) = 0, \quad y'(0) = 1.$$

$$5.46. y'' + 2y' + 5y = 4t, \quad y(0) = 0, \quad y'(0) = 0.$$

$$5.47. y'' - 4y' + 4y = t^3 e^{2t}, \quad y(0) = 0, \quad y'(0) = 0.$$

$$5.48. y'' + 4y = \begin{cases} 1, & 0 \leq t < 1, \\ 0, & t \geq 1, \end{cases} \quad y(0) = 0, \quad y'(0) = -1.$$

$$5.49. y'' - 5y' + 6y = \begin{cases} t, & 0 \leq t < 1 \\ 0, & t \geq 1 \end{cases}, \quad y(0) = 0, \quad y'(0) = 1.$$

$$5.50. y'' + 4y' + 3y = \begin{cases} 4e^{1-t}, & 0 \leq t < 1 \\ 4, & t \geq 1 \end{cases}, \quad y(0) = 0, \quad y'(0) = 0.$$

$$*5.51. y'' + 4y' = u(t-1), \quad y(0) = 0, \quad y'(0) = 0.$$

$$5.52. y'' + 3y' + 2y = 1 - u(t-1), \quad y(0) = 0, \quad y'(0) = 1.$$

$$5.53. y'' - y = \sin t + \delta(t - \pi/2), \quad y(0) = 3.5, \quad y'(0) = -3.5.$$

$$5.54. y'' + 5y' + 6y = u(t-1) + \delta(t-2), \quad y(0) = 0, \quad y'(0) = 1.$$

Résoudre les équations intégrales par transformation de Laplace et tracer les solutions.

$$5.55. y(t) = 1 + \int_0^t y(\tau) d\tau.$$

$$5.56. y(t) = \sin t + \int_0^t y(\tau) \sin(t - \tau) d\tau.$$

$$5.57. y(t) = \cos 3t + 2 \int_0^t y(\tau) \cos 3(t - \tau) d\tau.$$

$$5.58. y(t) = t + e^t + \int_0^t y(\tau) \cosh(t - \tau) d\tau.$$

$$5.59. y(t) = t e^t + 2 e^t \int_0^t e^{-\tau} y(\tau) d\tau.$$

Tracer 3 périodes des fonctions 2π -périodiques et trouver leurs transformées de Laplace.

$$5.60. f(t) = \pi - t, \quad 0 < t < 2\pi.$$

$$5.61. f(t) = 4\pi^2 - t^2, \quad 0 < t < 2\pi.$$

$$5.62. f(t) = e^{-t}, \quad 0 < t < 2\pi.$$

$$5.63. f(t) = \begin{cases} t, & \text{si } 0 < t < \pi, \\ \pi - t, & \text{if } \pi < t < 2\pi. \end{cases}$$

$$5.64. f(t) = \begin{cases} 0, & \text{si } 0 < t < \pi, \\ t - \pi, & \text{si } \pi < t < 2\pi. \end{cases}$$

Exercices pour le chapitre 6

Trouver l'intervalle de convergence des séries et de leur 1ère dérivée terme à terme.

$$6.1. \sum_{n=1}^{\infty} \frac{(-1)^n}{2n+1} x^n.$$

$$6.2. \sum_{n=1}^{\infty} \frac{2^n}{n3^{n+3}} x^n.$$

$$6.3. \sum_{n=2}^{\infty} \frac{\ln n}{n} x^n.$$

$$6.4. \sum_{n=1}^{\infty} \frac{1}{n^2 + 1} (x + 1)^n.$$

$$6.5. \sum_{n=3}^{\infty} \frac{n(n-1)(n-2)}{4^n} x^n.$$

$$6.6. \sum_{n=0}^{\infty} \frac{(-1)^n}{k^n} x^{2n}.$$

$$6.7. \sum_{n=0}^{\infty} \frac{(-1)^n}{k^n} x^{3n}.$$

$$6.8. \sum_{n=1}^{\infty} \frac{(4n)!}{(n!)^4} x^n.$$

Trouver les solutions séries des équations différentielles.

$$6.9. y'' - 3y' + 2y = 0.$$

$$6.10. (1 - x^2)y'' - 2xy' + 2y = 0.$$

$$6.11. y'' + x^2y' + xy = 0.$$

$$*6.12. y'' - xy' - y = 0.$$

$$6.13. (x^2 - 1)y'' + 4xy' + 2y = 0.$$

$$6.14. (1 - x)y'' - y' + xy = 0.$$

$$6.15. y'' - 4xy' + (4x^2 - 2)y = 0.$$

$$6.16. y'' - 2(x - 1)y' + 2y = 0.$$

6.17. Montrer que la substitution $x = \cos \theta$ transforme l'équation

$$\sin \theta \frac{d^2 y}{d\theta^2} + \cos \theta \frac{dy}{d\theta} + n(n + 1)(\sin \theta)y = 0$$

en l'équation de Legendre.

6.18. Dériver la formule de Rodrigues (6.10).

6.19. Dériver la fonction génératrice (6.11).

6.20. Soit A_1 et A_2 deux points de l'espace (V. la fig. 12.1). Dériver la formule, importante en théorie du potentiel,

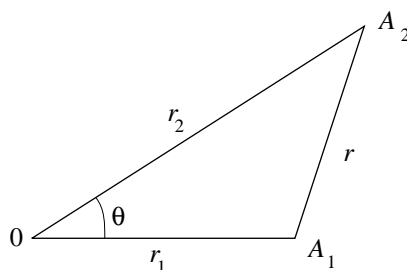
$$\frac{1}{r} = \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta}} = \frac{1}{r_2} \sum_{m=0}^{\infty} P_m(\cos \theta) \left(\frac{r_1}{r_2}\right)^m$$

au moyen de (6.9).

6.21. Dériver la récurrence de Bonnet,

$$(n + 1)P_{n+1}(x) = (2n + 1)xP_n(x) - nP_{n-1}(x), \quad n = 1, 2, \dots \quad (\text{a})$$

(Indication. Dériver la fonction génératrice (6.11) par rapport à t , substituer (6.11) dans la dérivée et comparer les coefficients de t^n .)

FIGURE 12.1. Distance r du point A_1 au point A_2 .

6.22. Comparer la valeur de $P_4(0.7)$ obtenue par la récurrence à 3 points (a) de l'exercice précédent avec la valeur du polynôme $P_4(x)$ en $x = 0.7$.

6.23. Pour deux entiers m et n tels que $0 \leq m \leq n$, posons

$$p_n^m(x) = \frac{d^m}{dx^m} P_n(x).$$

Montrer que $p_n^m(x)$ est solution de l'équation différentielle

$$(1 - x^2)y'' + 2(m + 1)xy' + (n - m)(n + m + 1)y = 0.$$

Représenter les polynômes suivants au moyen des polynômes de Legendre $P_0(x), P_1(x), \dots$

6.24. $p(x) = 5x^3 + 4x^2 + 3x + 2, \quad -1 \leq x \leq 1.$

6.25. $p(x) = 10x^3 + 4x^2 + 6x + 1, \quad -1 \leq x \leq 1.$

6.26. $p(x) = x^3 - 2x^2 + 4x + 1, \quad -1 \leq x \leq 2.$

Trouver les 3 premiers coefficients du développement de Fourier-Legendre des fonctions $f(x)$ et tracer $f(x)$ et les approximations de Fourier-Legendre sur le même repère.

6.27. $f(x) = e^x, \quad -1 < x < 1.$

6.28. $f(x) = e^{2x}, \quad -1 < x < 1.$

6.29. $f(x) = \begin{cases} 0 & -1 < x < 0, \\ 1 & 0 < x < 1. \end{cases}$

6.30. Intégrer numériquement

$$I = \int_{-1}^1 (5x^5 + 4x^4 + 3x^3 + 2x^2 + x + 1) dx,$$

par la quadrature gaussienne à 3 points. Trouver aussi la valeur exacte de I et calculer l'erreur dans la valeur numérique.

***6.31.** Évaluer

$$I = \int_{0.2}^{1.5} e^{-x^2} dx,$$

par la quadrature gaussienne à 3 points.

6.32. Évaluer

$$I = \int_{0.3}^{1.7} e^{-x^2} dx,$$

par la quadrature gaussienne à 3 points.

6.33. Dériver la quadrature gaussienne à 4 points.

6.34. Obtenir $P_4(x)$ par la formule de Bonnet de l'exercice 6.21 ou autrement.

6.35. Trouver les zéros de $P_4(x)$ sous formes de radicals.

Indication: Substituer $t = x^2$ dans le polynôme $P_4(x)$ du 4ème degré et résoudre l'équation du second degré.

6.36. Obtenir $P_5(x)$ par la formule de Bonnet de l'exercice 6.21 ou autrement.

6.37. Trouver les zéros de $P_5(x)$ sous formes de radicals.

Indication: Écrire $P_5(x) = xQ_4(x)$. Puis substituer $t = x^2$ dans le polynôme $Q_4(x)$ du 4ème degré et résoudre l'équation du second degré.

Exercices pour les méthodes numériques

Les solutions des exercices marqués d'un astérisque se trouvent au chapitre 14.
Les angles sont toujours en radians.

Exercices pour le chapitre 7

7.1. Calculer x_3 par dichotomie pour $f(x) = \sqrt{x} - \cos x$ sur $[0, 1]$. Les angles sont en radians.

7.2. Calculer x_3 par dichotomie pour

$$f(x) = 3(x+1)(x-1/2)(x-1)$$

sur les intervalles:

$$[-2, 1.5], \quad [-1.25, 2.5].$$

7.3. Calculer par dichotomie la solution à 10^{-3} près pour $f(x) = x - \tan x$ sur $[4, 4.5]$. Les angles sont en radians.

7.4. Faire 4 itérations pour calculer $\sqrt{3}$ par dichotomie. Démarrer en $a_0 = 1$ et $b_0 = 2$. Combien faut-il d'itérations pour avoir une réponse à 4 décimales près ?
[Indication: Considérer la fonction $f(x) = x^2 - 3$.]

7.5. Montrer que la récurrence de point fixe

$$x_{n+1} = \sqrt{2x_n + 3}$$

pour $f(x) = x^2 - 2x - 3 = 0$ converge sur l'intervalle $[2, 4]$.

7.6. Calculer, à 10^{-2} près, la solution de $f(x) = x^3 - x - 1 = 0$ par une récurrence de point fixe sur $[1, 2]$, autre que celle de Newton. Démarrer en $x_0 = 1$.

7.7. Calculer $\sqrt{3}$ à 4 décimales près par la récurrence de Newton. Démarrer en $x_0 = 1$. Comparer votre résultat et le nombre d'itérations avec ceux de l'exercice 7.4.

7.8. Itérer 5 fois la récurrence de point fixe $x_{n+1} = g(x_n)$ avec $g(x) = \cos(x-1)$. Démarrer en $x_0 = 2$. Utiliser au moins 6 décimales. Trouver l'ordre de convergence de la méthode. Les angles sont en radians.

7.9. Itérer 5 fois la récurrence de point fixe $x_{n+1} = g(x_n)$ avec $g(x) = 1 + \sin^2 x$. Démarrer en $x_0 = 1$. Utiliser au moins 6 décimales. Trouver l'ordre de convergence de la méthode. Les angles sont en radians.

7.10. Tracer la fonction $f(x) = 2x - \tan x$ et calculer une racine de $f(x) = 0$ à 6 décimales près par la méthode de Newton. Démarrer en $x_0 = 1$. Trouver l'ordre de convergence de la méthode. Les angles sont en radians.

***7.11.** Tracer la fonction $f(x) = e^{-x} - \tan x$ et calculer une racine de $f(x) = 0$ à 6 décimales près par la méthode de Newton. Démarrer en $x_0 = 1$. Trouver l'ordre de convergence de la méthode. Les angles sont en radians.

7.12 Calculer un zéro de la fonction $f(x) = 2x - \tan x$ (V. exercice 7.10) par la méthode de la sécante. Démarrer en $x_0 = 1$ et $x_1 = 0.5$. Trouver l'ordre de convergence de la méthode.

7.13. Refaire l'exercice 7.12 par la méthode de la position fausse. Trouver l'ordre de convergence de la méthode.

7.14. Refaire l'exercice 7.11 par la méthode de la sécante. Démarrer en $x_0 = 1$ et $x_1 = 0.5$. Trouver l'ordre de convergence de la méthode.

7.15. Refaire l'exercice 7.14 par la méthode de la position fausse. Trouver l'ordre de convergence de la méthode.

7.16. Soit la récurrence de point fixe de l'exercice 7.5:

$$x_{n+1} = \sqrt{2x_n + 3}.$$

Compléter le tableau:

| n | x_n | Δx_n | $\Delta^2 x_n$ |
|-----|------------------------------|----------------------|----------------------|
| 1 | $x_1 = 4.000$ | | |
| 2 | $x_2 =$ <input type="text"/> | <input type="text"/> | <input type="text"/> |
| 3 | $x_3 =$ <input type="text"/> | <input type="text"/> | |

Accélérer la convergence par la méthode d'Aitken.

$$a_1 = x_1 - \frac{(\Delta x_1)^2}{\Delta^2 x_1} = \text{}$$

7.17. Appliquer la méthode de Steffensen au résultat de l'exercice 7.9. Trouver l'ordre de convergence de la méthode.

7.18. Calculer les trois zéros de

$$f(x) = x^3 + 3x^2 - 1$$

par la méthode de Müller.

7.19. Calculer les quatre zéros de

$$f(x) = x^4 + 2x^2 - x - 3$$

par la méthode de Müller.

7.20. Tracer la fonction $f(x) = x - \tan x$. Trouver la multiplicité du zéro en $x = 0$. Calculer la racine de $f(x) = 0$ en $x = 0$ à 6 décimales près par la méthode de Newton modifiée qui tient compte de la multiplicité de la racine. Démarrer en $x_0 = 1$. Trouver l'ordre de convergence de la méthode de Newton modifiée employée.

***7.21.** Tracer la fonction $f(x) = x - \tan x$. Trouver la multiplicité du zéro en $x = 0$. Calculer la racine de $f(x) = 0$ en $x = 0$ à 6 décimales près par la méthode de sécante. Démarrer en $x_0 = 1$ et $x_1 = 0.5$. Trouver l'ordre de convergence de la méthode.

Exercices pour le chapitre 8

8.1. Construire les polynômes de Lagrange de degré 1 et 2 qui interpolent la fonction $f(x) = \ln(x + 1)$ en 2 et en 3 des points $x_0 = 0$, $x_1 = 0.6$ et $x_2 = 0.9$. Évaluer les polynômes en $x = 0.45$ et trouver l'erreur actuelle.

8.2. Soit les données

$$f(8.1) = 16.94410, \quad f(8.3) = 17.56492, \quad f(8.6) = 18.50515, \quad f(8.7) = 18.82091.$$

Interpoler $f(8.4)$ par des polynômes de Lagrange de degré 1, 2 et 3.

8.3. Construire le polynôme de Lagrange qui interpole la fonction $f(x) = e^{2x} \cos 3x$ en $x_0 = 0$, $x_1 = 0.3$ et $x_2 = 0.6$.

***8.4.** Soit trois points

$$(0.1, 1.0100502), \quad (0.2, 1.04081077), \quad (0.4, 1.1735109)$$

sur le graphe d'une fonction $f(x)$. Employer ces points pour estimer $f(0.3)$.

8.5. Compléter le tableau de différences divisées:

| i | x_i | $f[x_i]$ | $f[x_i, x_{i+1}]$ | $f[x_i, x_{i+1}, x_{i+2}]$ | $f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$ |
|-----|-------|----------|-------------------|----------------------------|-------------------------------------|
| 0 | 3.2 | 22.0 | | | |
| 1 | 2.7 | 17.8 | 8.400 | | |
| 2 | 1.0 | 14.2 | | 2.856 | |
| 3 | 4.8 | 38.3 | | | -0.528 |
| 4 | 5.6 | 5.17 | | | |

Construire le polynôme de degré 3 qui interpole les données aux 4 points de $x_0 = 3.2$ à $x_3 = 4.8$.

8.6. Construire le polynôme de Newton aux différences divisées de degré 3 qui interpole les données

$$(-1, 2), \quad (0, 0), \quad (1.5, -1), \quad (2, 4).$$

Tracer les données et le polynôme sur le même repère.

8.7. Répéter l'exercice 8.1 avec des polynômes de Newton aux différences divisées.

8.8. Répéter l'exercice 8.2 avec des polynômes de Newton aux différences divisées.

8.9. Interpoler les données

$$(-1, 2), \quad (0, 0), \quad (1, -1), \quad (2, 4),$$

par un polynôme de Gregory–Newton de degré 3.

8.10. Interpoler les données

$$(-1, 3), \quad (0, 1), \quad (1, 0), \quad (2, 5),$$

par un polynôme de Gregory–Newton de degré 3.

8.11. Approcher $f(0.05)$ par un polynôme de Gregory–Newton prograde (avant) de degré 4 qui interpole les données:

| x | 0.0 | 0.2 | 0.4 | 0.6 | 0.8 |
|--------|---------|---------|---------|---------|---------|
| $f(x)$ | 1.00000 | 1.22140 | 1.49182 | 1.82212 | 2.22554 |

***8.12.** Répéter l'exercice 8.11 avec un polynôme de Gregory–Newton rétrograde (arrière) de degré 4.

8.13. Construire un polynôme d'Hermite de degré 3 qui interpole les données:

| x | $f(x)$ | $f'(x)$ |
|-----|----------|----------|
| 8.3 | 17.56492 | 3.116256 |
| 8.6 | 18.50515 | 3.151762 |

Exercices pour le chapitre 9

9.1. Soit les formules de dérivation numérique:

$$(DN.4) \quad f'(x_0) = \frac{1}{2h}[-3f(x_0) + 4f(x_0 + h) - f(x_0 + 2h)] + \frac{h^2}{3}f^{(3)}(\xi),$$

$$(DN.5) \quad f'(x_0) = \frac{1}{2h}[f(x_0 + h) - f(x_0 - h)] - \frac{h^2}{6}f^{(3)}(\xi),$$

$$(DN.6) \quad f'(x_0) = \frac{1}{12h}[f(x_0 - 2h) - 8f(x_0 - h) + 8f(x_0 + h) - f(x_0 + 2h)] + \frac{h^4}{30}f^{(5)}(\xi),$$

$$(DN.7) \quad f'(x_0) = \frac{1}{12h}[-25f(x_0) + 48f(x_0 + h) - 36f(x_0 + 2h) + 16f(x_0 + 3h) - 3f(x_0 + 4h) + \frac{h^4}{5}f^{(5)}(\xi),$$

et le tableau $\{x_n, f(x_n)\}$:

`x = 1:0.1:1.8; format long; table = [x', (cosh(x)-sinh(x))']`

`table =`

| | |
|-------------------|------------------|
| 1.000000000000000 | 0.36787944117144 |
| 1.100000000000000 | 0.33287108369808 |
| 1.200000000000000 | 0.30119421191220 |
| 1.300000000000000 | 0.27253179303401 |
| 1.400000000000000 | 0.24659696394161 |
| 1.500000000000000 | 0.22313016014843 |
| 1.600000000000000 | 0.20189651799466 |
| 1.700000000000000 | 0.18268352405273 |
| 1.800000000000000 | 0.1652988822159 |

Pour les formules (DN.4)–(DN.7) et $h = 0.1$,

(a) calculer les valeurs numériques

$$\text{dfn} = f'(1.2)$$

(en omettant l'erreur de méthode);

(b) calculer la valeur exacte

$$\text{dfe} = f'(1.2)$$

de la fonction

$$f(x) = \cosh x - \sinh x;$$

(c) calculer l'erreur

$$\varepsilon = \text{dfn} - \text{dfe};$$

(d) vérifier que la valeur de $|\varepsilon|$ est bornée par le module de l'erreur de méthode.

9.2. Soit $f(x) = x^2 e^{-x}$. Utiliser l'extrapolation de Richardson avec $h = 0.4$, $h/2$ et $h/4$ pour affiner la valeur de $f'(1.4)$ obtenue par la différence centrée (DN.5).

9.3. Évaluer $\int_0^1 \frac{dx}{1+x}$ par la méthode des trapèzes (composée) avec $n = 10$.

9.4. Évaluer $\int_0^1 \frac{dx}{1+x}$ par la méthode de Simpson (composée) avec $n = 2m = 10$.

9.5. Évaluer $\int_0^1 \frac{dx}{1+2x^2}$ par la méthode des trapèzes (composée) avec $n = 10$.

9.6. Évaluer $\int_0^1 \frac{dx}{1+2x^2}$ par la méthode de Simpson (composée) avec $n = 2m = 10$.

9.7. Approcher l'intégrale $\int_0^1 \frac{dx}{1+x^3}$ à 10^{-4} près par la méthode des trapèzes (composée).

9.8. Approcher l'intégrale $\int_0^1 \frac{dx}{1+x^3}$ à 10^{-6} près par la méthode de Simpson (composée).

9.9. Déterminer les valeurs de h et n pour approcher l'intégrale

$$\int_1^3 \ln x \, dx$$

à 10^{-3} près par les méthodes (composées) des trapèzes, de Simpson et des points milieux.

9.10. Répéter l'exercice 9.9 avec

$$\int_0^2 \frac{1}{x+4} \, dx$$

à 10^{-5} près.

9.11. Employer l'intégration de Romberg pour calculer $R_{3,3}$ pour l'intégrale

$$\int_1^{1.5} x^2 \ln x \, dx.$$

9.12. Employer l'intégration de Romberg pour calculer $R_{3,3}$ pour l'intégrale

$$\int_1^{1.6} \frac{2x}{x^2 - 4} dx.$$

9.13. Appliquer l'intégration de Romberg sur l'intégrale

$$\int_0^1 x^{1/3} dx$$

jusqu'à ce que $|R_{n-1,n-1} - R_{n,n}| \leq 10^{-4}$.

Exercices pour le chapitre 10

Approcher à 4 décimales près la solution des problèmes à valeur initiales sur $0 \leq x \leq 1$ par la méthode d'Euler avec $h = 0.1$ et tracer la solution numérique.

10.1. $y' = e^{-y} - y + 1, \quad y(0) = 1.$

10.2. $y' = x + \sin y, \quad y(0) = 0.$

***10.3.** $y' = x + \cos y, \quad y(0) = 0.$

10.4. $y' = x^2 + y^2, \quad y(0) = 1.$

10.5. $y' = 1 + y^2, \quad y(0) = 0.$

Approcher à 4 décimales près la solution des problèmes à valeur initiale sur $0 \leq x \leq 1$ par la méthode d'Euler améliorée avec $h = 0.1$ et tracer la solution numérique.

10.6. $y' = e^{-y} - y + 1, \quad y(0) = 1.$

10.7. $y' = x + \sin y, \quad y(0) = 0.$

***10.8.** $y' = x + \cos y, \quad y(0) = 0.$

10.9. $y' = x^2 + y^2, \quad y(0) = 1.$

10.10. $y' = 1 + y^2, \quad y(0) = 0.$

Approcher à 4 décimales près la solution des problèmes à valeur initiale sur $0 \leq x \leq 1$ par la méthode de Runge-Kutta d'ordre 4 avec $h = 0.1$ et tracer la solution numérique.

10.11. $y' = x^2 + y^2, \quad y(0) = 1.$

10.12. $y' = x + \sin y, \quad y(0) = 0.$

***10.13.** $y' = x + \cos y, \quad y(0) = 0.$

10.14. $y' = e^{-y}, \quad y(0) = 0.$

10.15. $y' = y^2 + 2y - x, \quad y(0) = 0.$

Approcher à 6 décimales près la solution des problèmes à valeur initiale sur $0 \leq x \leq 1$ par la paire de Matlab `ode23` d'ordre 3 avec $h = 0.1$ et estimer l'erreur locale de méthode par la formule donnée.

10.16. $y' = x^2 + 2y^2, \quad y(0) = 1.$

$$10.17. y' = x + 2 \sin y, \quad y(0) = 0.$$

$$10.18. y' = x + 2 \cos y, \quad y(0) = 0.$$

$$10.19. y' = e^{-y}, \quad y(0) = 0.$$

$$10.20. y' = y^2 + 2y - x, \quad y(0) = 0.$$

Approcher à 6 décimales près la solution des problèmes à valeur initiale sur $0 \leq x \leq 1$ par le prédicteur-correcteur d'Adams–Bashforth–Moulton d'ordre 3 avec $h = 0.1$, estimer l'erreur locale de méthode en $x = 0.5$ et tracer la solution numérique.

$$10.21. y' = x + \sin y, \quad y(0) = 0.$$

$$10.22. y' = x + \cos y, \quad y(0) = 0.$$

$$10.23. y' = y^2 - y + 1, \quad y(0) = 0.$$

Approcher à 6 décimales près la solution des problèmes à valeur initiale sur $0 \leq x \leq 1$ par le prédicteur-correcteur d'Adams–Bashforth–Moulton d'ordre 4 avec $h = 0.1$, estimer l'erreur locale de méthode en $x = 0.5$ et tracer la solution numérique.

$$10.24. y' = x + \sin y, \quad y(0) = 0.$$

$$*10.25. y' = x + \cos y, \quad y(0) = 0.$$

$$10.26. y' = y^2 - y + 1, \quad y(0) = 0.$$

Exercices pour le chapitre 11

Résoudre les systèmes par décomposition LU sans pivotage.

11.1.

$$\begin{array}{rclcl} 2x_1 & + & 2x_2 & + & 2x_3 & = & 4 \\ -x_1 & + & 2x_2 & - & 3x_3 & = & 32 \\ 3x_1 & & & - & 4x_3 & = & 17 \end{array}$$

*11.2.

$$\begin{array}{rclcl} x_1 & + & x_2 & + & x_3 & = & 5 \\ x_1 & + & 2x_2 & + & 2x_3 & = & 6 \\ x_1 & + & 2x_2 & + & 3x_3 & = & 8 \end{array}$$

Résoudre les systèmes par décomposition LU avec pivotage.

11.3.

$$\begin{array}{rclcl} 2x_1 & - & x_2 & + & 5x_3 & = & 4 \\ -6x_1 & + & 3x_2 & - & 9x_3 & = & -6 \\ 4x_1 & - & 3x_2 & & & = & -2 \end{array}$$

*11.4.

$$\begin{array}{rclcl} 3x_1 & + & 9x_2 & + & 6x_3 & = & 23 \\ 18x_1 & + & 48x_2 & - & 39x_3 & = & 136 \\ 9x_1 & - & 27x_2 & + & 42x_3 & = & 45 \end{array}$$

11.5. Étalonner le 1er membre de chaque équation dans la norme l_∞ puis résoudre le système étalonné par décomposition LU avec pivotage.

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 3.8 \\ 4x_1 + 3x_2 - x_3 &= -5.7 \\ 5x_1 + 10x_2 + 3x_3 &= 2.8 \end{aligned}$$

***11.6.** Trouver l'inverse de la transformation gaussienne.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -a & 1 & 0 & 0 \\ -b & 0 & 1 & 0 \\ -c & 0 & 0 & 1 \end{bmatrix}.$$

***11.7.** Faire le produit des 3 transformations gaussiennes.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & 0 & 1 & 0 \\ c & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & d & 1 & 0 \\ 0 & e & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & f & 1 \end{bmatrix}.$$

11.8. Faire la décomposition de Cholesky des matrices.

$$A = \begin{bmatrix} 1 & -1 & -2 \\ -1 & 5 & -4 \\ -2 & -4 & 22 \end{bmatrix}, \quad B = \begin{bmatrix} 9 & 9 & 9 & 0 \\ 9 & 13 & 13 & -2 \\ 9 & 13 & 14 & -3 \\ 0 & -2 & -3 & 18 \end{bmatrix}.$$

Résoudre les systèmes par la décomposition de Cholesky.

11.9.

$$\begin{bmatrix} 16 & -4 & 4 \\ -4 & 10 & -1 \\ 4 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -12 \\ 3 \\ 1 \end{bmatrix}.$$

***11.10.**

$$\begin{bmatrix} 4 & 10 & 8 \\ 10 & 26 & 26 \\ 8 & 26 & 61 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 44 \\ 128 \\ 214 \end{bmatrix}.$$

Permuter les systèmes pour assurer la convergence du schéma de Gauss-Seidel, puis itérer ce schéma 3 fois avec les valeurs de démarrage $\mathbf{x}^{(0)}$.

***11.11.**

$$\begin{aligned} 6x_1 + x_2 - x_3 &= 3 & x_1^{(0)} &= 1 \\ -x_1 + x_2 + 7x_3 &= -17 & x_2^{(0)} &= 1 \\ x_1 + 5x_2 + x_3 &= 0 & x_3^{(0)} &= 1 \end{aligned} \quad \text{avec}$$

11.12.

$$\begin{aligned} -2x_1 + x_2 + 6x_3 &= 22 & x_1^{(0)} &= 1 \\ x_1 + 4x_2 + 2x_3 &= 13 & x_2^{(0)} &= 1 \\ 7x_1 + 2x_2 - x_3 &= -6 & x_3^{(0)} &= 1 \end{aligned} \quad \text{avec}$$

11.13. Approcher les données

$$(0.9, 10), \quad (0.5, 5), \quad (1.6, 15), \quad (2.1, 20),$$

au sens des moindres carrés par la droite (s, F) où s est l'étirement d'un ressort de force de restauration F , et estimer la constante de restauration $k = F/s$. ($F = ks$ selon la loi de Hooke).

***11.14.** Approcher les données

$$(-1, 2), \quad (0, 0), \quad (1, 1), \quad (2, 2)$$

au sens des moindres carrés par une parabole.

11.15. Approcher les données

$$(0, 3.7), \quad (1, 3.0), \quad (2, 2.4), \quad (3, 1.8)$$

au sens des moindres carrés par la fonction $f(x) = a_0 + a_1 \cos x$. Note: x en radians.

11.16. Approcher les données

| | | | | | | | |
|-------|----------|------------|---|-----------|-----------|-----------|-----|
| x_i | -1 | -0.5 | 0 | 0.25 | 0.5 | 0.75 | 1 |
| y_i | e^{-1} | $e^{-1/2}$ | 1 | $e^{1/4}$ | $e^{1/2}$ | $e^{3/4}$ | e |

au sens des moindres carrés par la fonction

$$f(x) = a_0 P_0(x) + a_1 P_1(x) + a_2 P_2(x),$$

où P_0 , P_1 et P_2 sont les polynômes de Legendre de degré 0, 1 et 2. Tracer $f(x)$ et $g(x) = e^x$ sur le même repère.

Tracer les disques de Gershgorin qui contiennent les valeurs propres des matrices d'après le théorème 11.3.

11.17.

$$\begin{bmatrix} -i & 0.1 + 0.1i & 0.5i \\ 0.3i & 2 & 0.3 \\ 0.2 & 0.3 + 0.4i & i \end{bmatrix}.$$

***11.18.**

$$\begin{bmatrix} -2 & 1/2 & i/2 \\ 1/2 & 0 & i/2 \\ -i/2 & -i/2 & 2 \end{bmatrix}.$$

11.19. Calculer la norme l_1 de la matrice de l'exercice 11.17 et la norme l_∞ de la matrice de l'exercice 11.18.

Trouver la plus grande valeur propre en module et le vecteur propre associé des matrices par 3 itérations de la méthode de puissance

11.20. $\begin{bmatrix} 10 & 4 \\ 4 & 2 \end{bmatrix}$ avec $\mathbf{x}^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

11.21. $\begin{bmatrix} 3 & 2 & 3 \\ 2 & 6 & 6 \\ 3 & 6 & 3 \end{bmatrix}$ avec $\mathbf{x}^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Solutions des exercices marqués d'un astérisque

Solutions des exercices du chapitre premier

Ex. 1.3. Résoudre $(1 + x^2)y' = \cos^2 y$.

SOLUTION. On sépare les variables,

$$\frac{dy}{\cos^2 y} = \frac{dx}{1 + x^2},$$

et l'on intègre,

$$\int \frac{dy}{\cos^2 y} = \int \frac{dx}{1 + x^2} + c,$$

ce qui donne

$$\tan y = \arctan x + c.$$

La solution générale est donc $y = \arctan(\arctan x + c)$. □

Ex. 1.11. Résoudre $xy' = y + \sqrt{y^2 - x^2}$.

SOLUTION. On récrit l'équation sous forme différentielle :

$$x \frac{dy}{dx} = y + \sqrt{y^2 - x^2} \quad \Rightarrow \quad (y + \sqrt{y^2 - x^2}) dx - x dy = 0.$$

Alors

$$M(x, y) = y + \sqrt{y^2 - x^2} \quad \text{et} \quad N(x, y) = -x,$$

sont homogènes de degré 1. Si l'on pose

$$y = ux \quad \text{et} \quad dy = u dx + x du$$

dans l'équation différentielle, on obtient

$$(ux + \sqrt{u^2 x^2 - x^2}) dx - x(u dx + x du) = 0,$$

ou

$$ux dx + x\sqrt{u^2 - 1} dx - xu dx - x^2 du = 0,$$

ou

$$x\sqrt{u^2 - 1} dx - x^2 du = 0,$$

ou

$$\frac{du}{\sqrt{u^2 - 1}} = \frac{dx}{x}.$$

Une intégration

$$\int \frac{du}{\sqrt{u^2 - 1}} = \int \frac{dx}{x} + c$$

donne

$$\operatorname{arccosh} u = \ln |x| + c \quad \Rightarrow \quad u = \cosh(\ln |x| + c).$$

La solution générale est donc $y = x \cosh(\ln |x| + c)$. □

Ex. 1.21. Résoudre le problème à valeur initiale :

$$(2xy - 3) dx + (x^2 + 4y) dy = 0, \quad y(1) = 2.$$

SOLUTION. On voit que l'équation différentielle est exacte :

$$\begin{aligned} M(x, y) &= 2xy - 3 \quad \text{et} \quad N(x, y) = x^2 + 4y, \\ M_y &= 2x \quad \text{et} \quad N_x = 2x, \quad \Rightarrow \quad M_y = N_x. \end{aligned}$$

On intègre l'équation par la méthode pratique :

$$\begin{aligned} u(x, y) &= \int M(x, y) dx + T(y) \\ &= \int (2xy - 3) dx + T(y) \\ &= x^2y - 3x + T(y). \end{aligned}$$

On trouve $T(y)$:

$$\begin{aligned} \frac{\partial u}{\partial y} &= \frac{\partial}{\partial y} (x^2y - 3x + T(y)) = x^2 + T'(y) = N(x, y) = x^2 + 4y, \\ T'(y) &= 4y \quad \Rightarrow \quad T(y) = 2y^2. \end{aligned}$$

Alors,

$$u(x, y) = x^2y - 3x + 2y^2,$$

et $u(x, y) = c$ donne la solution générale :

$$x^2y - 3x + 2y^2 = c.$$

La condition initiale $y(1) = 2$ détermine la constante c :

$$(1)^2(2) - 3(1) + 2(2)^2 = c \quad \Rightarrow \quad c = 7.$$

La solution (unique) est $x^2y - 3x + 2y^2 = 7$. □

Ex. 1.25. Trouver la solution générale de $(x + y^2) dx - 2xy dy = 0$.

SOLUTION. On voit que l'équation différentielle n'est pas exacte :

$$\begin{aligned} M(x, y) &= x + y^2 \quad \text{et} \quad N(x, y) = -2xy, \\ M_y &= 2y \quad \text{et} \quad N_x = -2y, \quad \Rightarrow \quad M_y \neq N_x. \end{aligned}$$

Cherchons un facteur d'intégration :

$$\frac{M_y - N_x}{N} = \frac{4y}{-2xy} = -\frac{2}{x},$$

une fonction de x seulement. Alors,

$$\mu(x) = e^{\int -\frac{2}{x} dx} = e^{-2 \ln x} = e^{\ln x^{-2}} = x^{-2}.$$

On multiplie l'équation différentielle par μ :

$$(x^{-1} + x^{-2}y^2) dx - 2x^{-1}y dy = 0.$$

On vérifie que cette équation est exacte :

$$\begin{aligned} M^*(x, y) &= x^{-1} + x^{-2}y^2 \quad \text{et} \quad N^*(x, y) = -2x^{-1}y, \\ M_y^* &= 2x^{-2}y \quad \text{et} \quad N_x^* = 2x^{-2}y, \quad \Rightarrow \quad M_y^* = N_x^*. \end{aligned}$$

On intègre l'équation par la méthode pratique :

$$\begin{aligned} u(x, y) &= \int N^*(x, y) dy + T(x) \\ &= \int -2x^{-1}y dy + T(x) \\ &= -x^{-1}y^2 + T(x). \end{aligned}$$

On trouve $T(x)$:

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial}{\partial x} (-x^{-1}y^2 + T(x)) = x^{-2}y^2 + T'(x) = M^*(x, y) = x^{-1} + x^{-2}y^2, \\ T'(x) &= x^{-1} \quad \Rightarrow \quad T(x) = \ln|x|. \end{aligned}$$

Alors,

$$u(x, y) = \ln|x| - x^{-1}y^2$$

et la solution générale $u(x, y) = c_1$ est

$$\ln|x| - x^{-1}y^2 = c_1,$$

ou $x^{-1}y^2 = c + \ln|x|$, ou $y^2 = cx + x \ln|x|$. □

Ex. 1.39. Résoudre le problème à valeur initiale : $y' + y \cos x = \cos x$, $y(0) = 1$.

SOLUTION. On a une équation différentielle linéaire du premier ordre de la forme $y' + f(x)y = r(x)$ où $f(x) = \cos x$ et $r(x) = \cos x$. Alors la solution générale est

$$\begin{aligned} y(x) &= e^{-\int \cos x dx} \left[\int e^{\int \cos x dx} \cos x dx + c \right] \\ &= e^{-\sin x} \left[\int e^{\sin x} \cos x dx + c \right] \\ &= e^{-\sin x} [e^{\sin x} + c] \\ &= 1 + ce^{-\sin x}. \end{aligned}$$

Enfin, la condition initiale détermine la constante c :

$$y(0) = 1 \quad \Rightarrow \quad 1 = 1 + ce^{-\sin 0} = 1 + c \quad \Rightarrow \quad c = 0,$$

et l'unique solution est $y(x) = 1$. □

Solutions des exercices du chapitre 2

Ex. 2.3. Résoudre l'équation différentielle $y'' - 9y' + 20y = 0$.

SOLUTION. L'équation caractéristique est

$$\lambda^2 - 9\lambda + 20 = (\lambda - 4)(\lambda - 5) = 0,$$

et la solution générale est $y(x) = c_1 e^{4x} + c_2 e^{5x}$. □

Ex. 2.13. Résoudre l'équation différentielle d'Euler-Cauchy $x^2 y'' + 3xy' - 3y = 0$.

SOLUTION. L'équation caractéristique est

$$m(m-1) + 3m - 3 = m^2 + 2m - 3 = (m+3)(m-1) = 0,$$

et la solution générale est $y(x) = c_1 x + c_2 x^{-3}$. □

Solutions des exercices du chapitre 3

Ex. 3.1. Résoudre l'équation différentielle à coefficients constants ; $y''' + 6y'' = 0$.

SOLUTION. L'équation caractéristique est

$$\lambda^3 + 6\lambda^2 = \lambda^2(\lambda + 6) = 0,$$

et la solution générale est $y(x) = c_1 + c_2x + c_3e^{-6x}$. \square

Ex. 3.9. Déterminer si les fonctions

$$y_1(x) = x, \quad y_2(x) = x^2, \quad y_3(x) = 2x - 5x^2$$

sont linéairement dépendantes ou indépendantes sur $-\infty < x < +\infty$.

SOLUTION. Puisque

$$y_3(x) = 2x - 5x^2 = 2y_1(x) - 5y_2(x),$$

les fonctions sont linéairement dépendantes. \square

Ex. 3.13. Montrer au moyen du wronskien que les fonctions

$$y_1(x) = e^x, \quad y_2(x) = e^{2x}, \quad y_3(x) = e^{-x}$$

sont linéairement indépendantes sur l'intervalle $-\infty < x < +\infty$.

SOLUTION. On évalue le wronskien de y_1, y_2, y_3 :

$$\begin{aligned} \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{vmatrix} &= \begin{vmatrix} e^x & e^{2x} & e^{-x} \\ e^x & 2e^{2x} & -e^{-x} \\ e^x & 4e^{2x} & e^{-x} \end{vmatrix} \\ &= (e^x)(e^{2x})(e^{-x}) \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & 4 & 1 \end{vmatrix} \\ &= e^{2x} \left[\begin{vmatrix} 2 & -1 \\ 4 & 1 \end{vmatrix} - \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} \right] \\ &= e^{2x} [(2+4) - (1+1) + (4-2)] \\ &= 6e^{2x} \neq 0 \end{aligned}$$

pour tout x . Puisque les trois fonctions admettent des dérivées continues d'ordre 3, par le corollaire 3.2 elles sont solutions d'une même équation différentielle; alors elles sont linéairement indépendantes sur $-\infty < x < +\infty$. \square

Ex. 3.25. Résoudre l'équation différentielle non homogène : $y'' - y' = e^x \sin x$.

SOLUTION. L'équation différentielle homogène, $y'' - y' = 0$, admet le polynôme caractéristique

$$\lambda^2 - \lambda = \lambda(\lambda - 1) = 0$$

et la solution générale

$$y_h(x) = c_1 + c_2 e^x.$$

Puisque la dimension de l'espace des dérivées du second membre $r(x) = e^x \sin x$ est 2, on peut procéder par coefficients indéterminés. Prenons

$$y_p(x) = a e^x \cos x + b e^x \sin x.$$

Alors

$$y_p'(x) = a e^x \cos x - a e^x \sin x + b e^x \sin x + b e^x \cos x$$

et

$$\begin{aligned} y_p''(x) &= a e^x \cos x - 2a e^x \sin x - a e^x \cos x + b e^x \sin x + 2b e^x \cos x - b e^x \sin x \\ &= -2a e^x \sin x + 2b e^x \cos x. \end{aligned}$$

Donc,

$$\begin{aligned} y_p''(x) - y_p'(x) &= -2a e^x \sin x + 2b e^x \cos x \\ &\quad - (a e^x \cos x - a e^x \sin x + b e^x \sin x + b e^x \cos x) \\ &= -a e^x \cos x - a e^x \sin x + b e^x \cos x - b e^x \sin x \\ &= (b - a) e^x \cos x + (-a - b) e^x \sin x \\ &= r(x) = e^x \sin x. \end{aligned}$$

Par identification des coefficients on a $b - a = 0$ et $-a - b = 1 \Rightarrow a = b = -1/2$. La solution particulière est donc

$$y_p(x) = -\frac{1}{2} e^x \cos x - \frac{1}{2} e^x \sin x$$

et la solution générale est

$$\begin{aligned} y(x) &= y_h(x) + y_p(x) \\ &= c_1 + c_2 e^x - \frac{1}{2} e^x \cos x - \frac{1}{2} e^x \sin x. \end{aligned}$$

□

Ex. 3.35. Résoudre l'équation différentielle non homogène

$$y'' + 3y' + 2y = \frac{1}{1 + e^x}.$$

SOLUTION. L'équation différentielle homogène, $y'' + 3y' + 2y = 0$, admet le polynôme caractéristique

$$\lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2) = 0$$

et la solution générale

$$y_h(x) = A y_1(x) + B y_2(x) = A e^{-x} + B e^{-2x}.$$

Le second membre, $r(x) = \frac{1}{1 + e^x}$, admet un nombre infini de dérivées indépendantes.

Donc par variation des paramètres on prend la solution particulière de l'équation différentielle homogène de la forme

$$y_p(x) = c_1(x)y_1(x) + c_2(x)y_2(x).$$

Le système à résoudre pour $c_1'(x)$ et $c_2'(x)$ est

$$\begin{bmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{bmatrix} \begin{bmatrix} c_1'(x) \\ c_2'(x) \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{1+e^x} \end{bmatrix},$$

c'est-à-dire

$$e^{-x} c_1' + e^{-2x} c_2' = 0, \tag{a}$$

$$-e^{-x} c_1' - 2e^{-2x} c_2' = \frac{1}{1 + e^x}. \tag{b}$$

Alors, (a)+(b) donne

$$-e^{-2x} c_2' = \frac{1}{1 + e^x}$$

ou

$$\begin{aligned} c_2' &= \frac{-e^{2x}}{1+e^x} = \frac{1-e^{2x}-1}{1+e^x} \\ &= \frac{(1+e^x)(1-e^x)-1}{1+e^x} = 1-e^x - \frac{1}{1+e^x}. \end{aligned}$$

Donc

$$c_2'(x) = 1 - e^x - \frac{e^{-x}}{e^{-x} + 1},$$

et

$$c_2(x) = x - e^x + \ln(e^{-x} + 1).$$

Alors, (a) implique

$$c_1' = -e^{-x}c_2' = \frac{e^x}{1+e^x} \Rightarrow c_1(x) = \ln(e^x + 1)$$

et la solution particulière est

$$\begin{aligned} y_p(x) &= c_1(x)y_1(x) + c_2(x)y_2(x) \\ &= \ln(e^x + 1)e^{-x} + [x - e^x + \ln(e^{-x} + 1)]e^{-2x} \\ &= e^{-x}\ln(e^x + 1) + xe^{-2x} - e^{-x} + e^{-2x}\ln(e^{-x} + 1). \end{aligned}$$

Puisque e^{-x} apparaît déjà dans $y_h(x)$, on peut éliminer ce terme et prendre

$$y_p(x) = e^{-x}\ln(e^x + 1) + xe^{-2x} + e^{-2x}\ln(e^{-x} + 1).$$

La solution générale est

$$y(x) = Ae^{-x} + Be^{-2x} + e^{-x}\ln(e^x + 1) + xe^{-2x} + e^{-2x}\ln(e^{-x} + 1). \quad \square$$

Solutions des exercices du chapitre 4

Ex. 4.3. Résoudre le système $\mathbf{y}' = \begin{bmatrix} -1 & 1 \\ 4 & -1 \end{bmatrix} \mathbf{y}$.

SOLUTION. On trouve les valeurs propres et les vecteurs propres de la matrice A du système :

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -1 - \lambda & 1 \\ 4 & -1 - \lambda \end{vmatrix} \\ &= (-1 - \lambda)^2 - 4 = \lambda^2 + 2\lambda - 3 = (\lambda + 3)(\lambda - 1) = 0; \end{aligned}$$

donc $\lambda_1 = -3$ and $\lambda_2 = 1$.

Pour $\lambda_1 = -3$,

$$(A - \lambda_1 I)\mathbf{u}_1 = \begin{bmatrix} -1 & 1 \\ 4 & -1 \end{bmatrix} \mathbf{u}_1 = \mathbf{0}, \Rightarrow \mathbf{u}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

Pour $\lambda_2 = 1$,

$$(A - \lambda_2 I)\mathbf{u}_2 = \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix} \mathbf{u}_2 = \mathbf{0}, \Rightarrow \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

La solution générale est

$$\mathbf{y}(x) = c_1 e^{-3x} \begin{bmatrix} 1 \\ -2 \end{bmatrix} + c_2 e^x \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \quad \square$$

Solutions des exercices du chapitre 5

Ex. 5.5. Trouver la transformée de Laplace de $f(t) = \cos^2 t = \frac{1}{2}[1 + \cos(2t)]$.

SOLUTION.

$$F(s) = \mathcal{L}\{\cos^2 t\} = \mathcal{L}\left\{\frac{1}{2}[1 + \cos(2t)]\right\} = \frac{1}{2}\left(\frac{1}{s} + \frac{s}{s^2 + 4}\right). \quad \square$$

Ex. 5.12. Trouver la transformée de Laplace de $f(t) = u(t-1)t^2$.

SOLUTION.

$$F(s) = \mathcal{L}\{u(t-1)t^2\} = e^{-s}\mathcal{L}\{t^2 + 2t + 1\} = e^{-s}\left(\frac{2}{s^3} + \frac{2}{s^2} + \frac{1}{s}\right),$$

où l'on a transformé $f(t-1) = t^2$ en $f(t) = (t+1)^2 = t^2 + 2t + 1$. □

Ex. 5.23. Trouver la transformée de Laplace inverse de $F(s) = \frac{2s+1}{s^2+5s+6}$.

SOLUTION. Par fraction simple :

$$F(s) = \frac{2s+1}{s^2+5s+6} = \frac{2s+1}{(s+3)(s+2)} = \frac{A}{s+3} + \frac{B}{s+2} = \frac{5}{s+3} - \frac{3}{s+2}$$

puisque

$$A(s+2) + B(s+3) = 2s+1 \quad \Rightarrow \quad \begin{cases} A+B=2, & A=5, \\ 2A+3B=1, & B=-3. \end{cases}$$

Alors,

$$f(t) = \mathcal{L}^{-1}\left\{\frac{5}{s+3} - \frac{3}{s+2}\right\} = 5e^{-3t} - 3e^{-2t}. \quad \square$$

Ex. 5.29. Trouver la transformée de Laplace inverse de $F(s) = \frac{e^{-3s}}{s^2(s-1)}$.

SOLUTION. Par fraction simple :

$$\begin{aligned} F(s) &= \frac{e^{-3s}}{s^2(s-1)} = e^{-3s}\left[\frac{As+B}{s^2} + \frac{C}{s-1}\right] \\ &= e^{-3s}\left[\frac{-s-1}{s^2} + \frac{1}{s-1}\right] = e^{-3s}\left[\frac{1}{s-1} - \frac{1}{s} - \frac{1}{s^2}\right] \end{aligned}$$

puisque

$$(As+B)(s-1) + Cs^2 = 1 \quad \Rightarrow \quad \begin{cases} A+C=0, & B=-1, \\ B-A=0, & A=-1, \\ -B=1, & C=1. \end{cases}$$

Alors,

$$\begin{aligned} f(t) &= \mathcal{L}^{-1}\left\{\frac{e^{-3s}}{s^2(s-1)}\right\} = u(t-3)g(t-3) \\ &= u(t-3)[e^{t-3} - 1 - (t-3)] = u(t-3)(e^{t-3} - t + 2). \end{aligned}$$

□

Ex. 5.36. Trouver la transformée de Laplace $f(t) = te^{-t}\cos t$.

SOLUTION.

$$\begin{aligned} F(s) &= \mathcal{L}\{t e^{-t} \cos t\} = -\frac{d}{ds} \mathcal{L}\{e^{-t} \cos t\} \\ &= -\frac{d}{ds} \left[\frac{s+1}{(s+1)^2 + 1} \right] = -\left[\frac{(1)((s+1)^2 + 1) - 2(s+1)(s+1)}{((s+1)^2 + 1)^2} \right] \\ &= \frac{(s+1)^2 - 1}{((s+1)^2 + 1)^2} = \frac{s^2 + 2s}{(s^2 + 2s + 2)^2}. \end{aligned}$$

□

Ex. 5.51. Résoudre le problème à valeur initiale

$$y'' + 4y' = u(t-1), \quad y(0) = 0, \quad y'(0) = 0,$$

par transformation de Laplace.

SOLUTION. Soit $Y(s) = \mathcal{L}\{y(t)\}$. La transformée du problème est

$$\begin{aligned} \mathcal{L}\{y''\} + 4\mathcal{L}\{y'\} &= \mathcal{L}\{u(t-1)\}, \\ s^2 Y(s) - sy(0) - y'(0) + 4(sY(s) - y(0)) &= \frac{e^{-s}}{s}, \\ (s^2 + 4s)Y(s) = \frac{e^{-s}}{s} &\Rightarrow Y(s) = \frac{e^{-s}}{s^2(s+4)}. \end{aligned}$$

Par fraction simple,

$$Y(s) = e^{-s} \left[\frac{As+B}{s^2} + \frac{C}{s+4} \right].$$

On chasse les dénominateurs :

$$(As+B)(s+4) + Cs^2 = 1 \Rightarrow \begin{cases} A+C=0, & B=1/4, \\ 4A+B=0, & A=-1/16 \\ 4B=1, & C=1/16. \end{cases}$$

Alors,

$$Y(s) = \frac{e^{-s}}{16} \left[\frac{-s+4}{s^2} + \frac{1}{s+4} \right] = \frac{e^{-s}}{16} \left[\frac{4}{s^2} - \frac{1}{s} + \frac{1}{s+4} \right].$$

La solution du problème est donc

$$\begin{aligned} y(t) &= \mathcal{L}^{-1}\{Y(s)\} = \mathcal{L}^{-1} \left\{ \frac{e^{-s}}{16} \left[\frac{4}{s^2} - \frac{1}{s} + \frac{1}{s+4} \right] \right\} \\ &= u(t-1)g(t-1) = \frac{1}{16} u(t-1) \left[4(t-1) - 1 + e^{-4(t-1)} \right] \\ &= \frac{1}{16} u(t-1) \left[4t - 5 + e^{-4(t-1)} \right]. \end{aligned}$$

□

Solutions des exercices du chapitre 6

Ex. 6.12. Trouver la solution série de l'équation différentielle $y'' - xy' - y = 0$.

SOLUTION. Soit

$$y(x) = \sum_{m=0}^{\infty} a_m x^m,$$

$$y'(x) = \sum_{m=0}^{\infty} m a_m x^{m-1} \quad \text{et} \quad y''(x) = \sum_{m=0}^{\infty} m(m-1) a_m x^{m-2}.$$

Alors,

$$\begin{aligned} y'' - xy' - y &= \sum_{m=0}^{\infty} m(m-1) a_m x^{m-2} - x \sum_{m=0}^{\infty} m a_m x^{m-1} - \sum_{m=0}^{\infty} a_m x^m \\ &= \sum_{m=0}^{\infty} (m+1)(m+2) a_{m+2} x^m - \sum_{m=0}^{\infty} m a_m x^m - \sum_{m=0}^{\infty} a_m x^m \\ &= \sum_{m=0}^{\infty} [(m+1)(m+2) a_{m+2} - m a_m x^m - a_m] x^m \\ &= 0, \quad \text{pour tout } x. \end{aligned}$$

Donc

$$(m+1)(m+2) a_{m+2} - (m+1) a_m = 0, \quad \text{pour tout } m,$$

et la relation de récurrence est

$$a_{m+2} = \frac{a_m}{m+2}.$$

Pour m pair,

$$a_2 = \frac{a_0}{2}, \quad a_4 = \frac{a_2}{4} = \frac{a_0}{2 \cdot 4}, \quad a_6 = \frac{a_4}{6} = \frac{a_0}{2 \cdot 4 \cdot 6}, \quad a_8 = \frac{a_6}{8} = \frac{a_0}{2 \cdot 4 \cdot 6 \cdot 8} = \frac{a_0}{2^4 4!},$$

d'où la formule

$$a_{2k} = \frac{a_0}{2^k k!}.$$

Pour m impair,

$$a_3 = \frac{a_1}{3}, \quad a_5 = \frac{a_3}{5} = \frac{a_1}{3 \cdot 5}, \quad a_7 = \frac{a_5}{7} = \frac{a_1}{5 \cdot 5 \cdot 7},$$

d'où la formule

$$a_{2k+1} = \frac{a_1}{1 \cdot 5 \cdot 5 \cdots (2k+1)} = \frac{a_1 2^k k!}{1 \cdot 3 \cdot 5 \cdots (2k+1) 2 \cdot 4 \cdot 6 \cdots (2k)} = \frac{2^k k! a_1}{(2k+1)!}.$$

La solution générale est

$$\begin{aligned} y(x) &= \sum_{m=0}^{\infty} a_m x^m = \sum_{k=0}^{\infty} a_{2k} x^{2k} + \sum_{k=0}^{\infty} a_{2k+1} x^{2k+1} \\ &= a_0 \sum_{k=0}^{\infty} \frac{x^{2k}}{k! 2^k} + a_1 \sum_{k=0}^{\infty} \frac{2^k k!}{(2k+1)!} x^{2k+1}. \end{aligned}$$

□

Ex. 6.31. Evaluer $I = \int_{0.2}^{1.5} e^{-x^2} dx$ par la quadrature gaussienne à trois points.

SOLUTION. On transforme les limites de l'intégrale de -1 à 1 par la substitution

$$x = \frac{1}{2} [0.2(1-t) + 1.5(t+1)] = \frac{1}{2} (1.3t + 1.7) = 0.65t + 0.85$$

et $dx = 0.65 dt$:

$$\begin{aligned} \int_{0.2}^{1.5} e^{-x^2} dx &= \int_{-1}^1 e^{-(0.65t+0.85)^2} (0.65) dt \\ &\approx 0.65 \sum_{j=1}^3 w_j f(t_j) \\ &= 0.65 \left[\frac{5}{9} e^{-(0.65(-0.7745966692)+0.85)^2} + \frac{8}{9} e^{-(0.65t(0)+0.85)^2} \right. \\ &\quad \left. + \frac{5}{9} e^{-(0.65(0.7745966692)+0.85)^2} \right] \\ &= 0.65 \left[\frac{5}{9} e^{-0.1200707} + \frac{8}{9} e^{-0.7225} + \frac{5}{9} e^{-1.8319292} \right] \\ &= 0.65(0.4926987 + 0.4315883 + 0.088946, 8) \\ &= 0.658602. \end{aligned}$$

□

Solutions des exercices du chapitre 7

Ex. 7.11. Tracer la fonction

$$f(x) = e^{-x} - \tan x$$

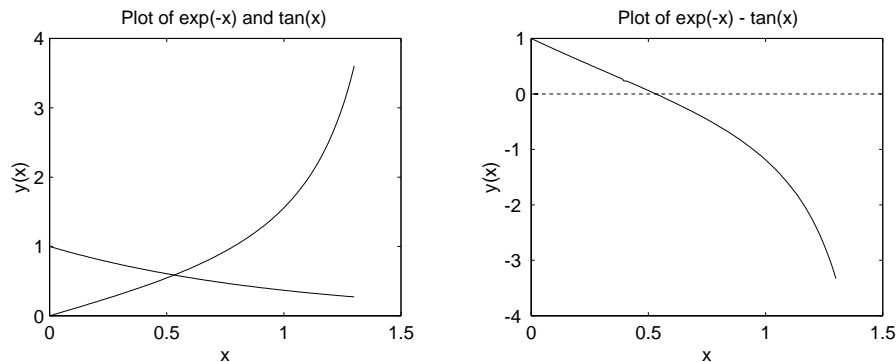
et calculer une racine de l'équation $f(x) = 0$ à six décimales près par la méthode de Newton avec $x_0 = 1$.

SOLUTION. On emploie le fichier M `newton7_11`

```
function f = newton7_11(x); % Exercice 7.11.
f = x - (exp(-x) - tan(x))/(-exp(-x) - sec(x)^2);
```

On itère par Newton et l'on vérifie la convergence à six décimales près.

```
>> xc = input('Entrer la valeur de depart:'); format long;
Entrer la valeur de depart:1
>> xc = newton7_11(xc)
xc = 0.68642146135728
>> xc = newton7_11(xc)
xc = 0.54113009740473
>> xc = newton7_11(xc)
xc = 0.53141608691193
>> xc = newton7_11(xc)
xc = 0.53139085681581
>> xc = newton7_11(xc)
xc = 0.53139085665216
```

FIGURE 14.1. Graphe de e^{-x} et $\tan x$ et de leur différence pour l'exercice 7.11.

Tous les chiffres de la dernière valeur x_c sont exacts. On voit que la convergence est d'ordre 2. Donc $x_c = 0.531391$ est la racine cherchée à six décimales près.

On trace les fonctions e^{-x} et $\tan x$ et leur différence. L'abscisse x du point d'intersection des deux fonctions est le zéro de leur différence.

```
x=0:0.01:1.3;
subplot(2,2,1); plot(x,exp(-x),x,tan(x));
title('Plot of exp(-x) and tan(x)'); xlabel('x'); ylabel('y(x)');
subplot(2,2,2); plot(x,exp(-x)-tan(x),x,0);
title('Plot of exp(-x)-tan(x)'); xlabel('x'); ylabel('y(x)');
print -deps Fig14_1
```

□

Ex. 7.21. Trouver la multiplicité du zéro en $x = 0$. Calculer la racine de $f(x) = 0$ en $x = 0$ à 6 décimales près par la méthode de la sécante. Démarrer en $x_0 = 1$ et $x_1 = 0.5$. Trouver l'ordre de convergence de la méthode.

SOLUTION. Puisque

$$f(0) = f'(0) = f''(0) = 0, \quad f'''(0) \neq 0,$$

$x = 0$ est un zéro triple. Le programme suivant exécute les calculs.

```
x0 = 1; x1 = 0.5; % valeurs de depart
x = zeros(20,1);
x(1) = x0; x(2) = x1;
for n = 3:20
    x(n) = x(n-1) - (x(n-1)-x(n-2)) ...
        / (x(n-1)-tan(x(n-1))-x(n-2)+tan(x(n-2)))*(x(n-1)-tan(x(n-1)));
end
dx = abs(diff(x));
p = 1; % verifier la convergence d'ordre 1
dxr = dx(2:19)./(dx(1:18).^p);
table = [[0:19]' x [0; dx] [0; 0; dxr]]
table =
    n          x_n      x_n - x_{n-1}    |x_n - x_{n-1}|
```

| | | | / $ x_{n-1} - x_{n-2} $ |
|----|------------------|------------------|-------------------------|
| 0 | 1.00000000000000 | | |
| 1 | 0.50000000000000 | 0.50000000000000 | |
| 2 | 0.45470356524435 | 0.04529643475565 | 0.09059286951131 |
| 3 | 0.32718945543123 | 0.12751410981312 | 2.81510256824784 |
| 4 | 0.25638399918811 | 0.07080545624312 | 0.55527546204022 |
| 5 | 0.19284144711319 | 0.06354255207491 | 0.89742451283310 |
| 6 | 0.14671560243705 | 0.04612584467614 | 0.72590481763723 |
| 7 | 0.11082587909404 | 0.03588972334302 | 0.77808273420254 |
| 8 | 0.08381567002072 | 0.02701020907332 | 0.75258894628889 |
| 9 | 0.06330169146740 | 0.02051397855331 | 0.75948980985777 |
| 10 | 0.04780894321090 | 0.01549274825651 | 0.75522884145761 |
| 11 | 0.03609714636358 | 0.01171179684732 | 0.75595347277403 |
| 12 | 0.02725293456160 | 0.00884421180198 | 0.75515413367179 |
| 13 | 0.02057409713542 | 0.00667883742618 | 0.75516479882196 |
| 14 | 0.01553163187404 | 0.00504246526138 | 0.75499146627099 |
| 15 | 0.01172476374403 | 0.00380686813002 | 0.75496169684658 |
| 16 | 0.00885088980844 | 0.00287387393559 | 0.75491817353192 |
| 17 | 0.00668139206035 | 0.00216949774809 | 0.75490358892216 |
| 18 | 0.00504365698583 | 0.00163773507452 | 0.75489134568691 |
| 19 | 0.00380735389990 | 0.00123630308593 | 0.75488588182657 |

On voit que $x_{19} = 0.0038$ est une solution approchée de la racine triple $x = 0$.
Puisque le rapport

$$\frac{|x_n - x_{n-1}|}{|x_{n-1} - x_{n-2}|} \rightarrow 0.75 \approx \text{constant}$$

quand n croît, on conclut que la méthode converge du premier ordre.

La convergence sur une racine triple est lente. En général, la méthode de la sécante peut ne pas converger sur une racine multiple. \square

Solutions des exercices du chapitre 8

Ex. 8.4. Les trois points

$$(0.1, 1.0100502), \quad (0.2, 1.04081077), \quad (0.4, 1.1735109)$$

sont sur le graphe d'une certaine fonction $f(x)$. Estimer $f(0.3)$ au moyen des trois points.

SOLUTION. On a les différences divisées :

$$f[0.1, 0.2] = \frac{1.04081077 - 1.0100502}{0.1} = 0.307606,$$

$$f[0.2, 0.4] = \frac{1.1735109 - 1.04081077}{0.2} = 0.663501$$

et

$$f[0.1, 0.2, 0.4] = \frac{0.663501 - 0.307606}{0.3} = 1.18632.$$

Alors,

$$p_2(x) = 1.0100502 + (x - 0.1) \times 0.307606 + (x - 0.1)(x - 0.2) \times 1.18632$$

et

$$p_2(0.3) = 1.0953. \quad \square$$

Ex. 8.12. Soit le tableau

| | | | | | |
|--------|---------|---------|---------|---------|---------|
| x | 0.0 | 0.2 | 0.4 | 0.6 | 0.8 |
| $f(x)$ | 1.00000 | 1.22140 | 1.49182 | 1.82212 | 2.22554 |

Approcher $f(0.65)$ par un polynôme de Gregory–Newton rétrograde de degré 4.

SOLUTION. Le programme

```
f = [1 1.2214 1.49182 1.82212 2.22554];
ddt = [f' [0 diff(f)]' [0 0 diff(f,2)]' [0 0 0 diff(f,3)]' ...
      [0 0 0 0 diff(f,4)]']
```

construit la table de différence :

| n | x_n | f_n | ∇f_n | $\nabla^2 f_n$ | $\nabla^3 f_n$ | $\nabla^4 f_n$ |
|-----|-------|--------|--------------|----------------|----------------|----------------|
| 0 | 0.0 | 1.0000 | | | | |
| | | | 0.22140 | | | |
| 1 | 0.2 | 1.2214 | | 0.04902 | | |
| | | | 0.27042 | | 0.01086 | |
| 2 | 0.4 | 1.4918 | | 0.05998 | | 0.00238 |
| | | | 0.33030 | | 0.01324 | |
| 3 | 0.6 | 1.8221 | | 0.07312 | | |
| | | | 0.40342 | | | |
| 4 | 0.8 | 2.2255 | | | | |

```
s = (0.65-0.80)/0.2 % la variable s
s = -0.7500
format long
p4 = ddt(5,1) + s*ddt(5,2) + s*(s+1)*ddt(5,3)/2 ...
      + s*(s+1)*(s+2)*ddt(5,4)/6 + s*(s+1)*(s+2)*(s+3)*ddt(5,5)/24
p4 = 1.91555051757812
```

□

Solutions des exercices du chapitre 10

On emploie le fichier M `exr10_25` pour les exercices 10.3, 10.8, 10.13 e 10.12 :

```
function yprime = exr10_25(x,y); % Exercices 10.3, 10.8, 10.13 et 10.25.
yprime = x+cos(y);
```

Ex. 10.3. Approcher la solution du problème à valeur initiale

$$y' = x + \cos y, \quad y(0) = 0,$$

sur $0 \leq x \leq 1$ à 4 décimale près par la méthode d'Euler avec $h = 0.1$ et tracer la solution numérique.

SOLUTION. **Solution numérique par Matlab.**— On résout l'équation différentielle par la méthode d'Euler :

```

clear
h = 0.1; x0= 0; xf= 1; y0 = 0;
n = ceil((xf-x0)/h); % nombre de pas
%
count = 2; print_time = 1; % pour l'écriture de donnees
x = x0; y = y0; % initialisation de x et de y
output1 = [0 x0 y0];
for i=1:n
    z = y + h*exr10_25(x,y);
    x = x + h;
    if count > print_time
        output1 = [output1; i x z];
        count = count - print_time;
    end
    y = z;
count = count + 1;
end
output1
save output1 % pour l'impression du graphe

```

La commande `output1` imprime les valeurs de n , x et y .

| n | x | y |
|-------------------|------------------|------------------|
| 0 | 0 | 0 |
| 1.00000000000000 | 0.10000000000000 | 0.10000000000000 |
| 2.00000000000000 | 0.20000000000000 | 0.20950041652780 |
| 3.00000000000000 | 0.30000000000000 | 0.32731391010682 |
| 4.00000000000000 | 0.40000000000000 | 0.45200484393704 |
| 5.00000000000000 | 0.50000000000000 | 0.58196216946658 |
| 6.00000000000000 | 0.60000000000000 | 0.71550074191996 |
| 7.00000000000000 | 0.70000000000000 | 0.85097722706339 |
| 8.00000000000000 | 0.80000000000000 | 0.98690209299587 |
| 9.00000000000000 | 0.90000000000000 | 1.12202980842386 |
| 10.00000000000000 | 1.00000000000000 | 1.25541526027779 |

□

Ex. 10.8. Soit le problème à valeur initiale

$$y' = x + \cos y, \quad y(0) = 0, \quad 0 \leq x \leq 1.$$

Approcher la solution à 4 décimales près par la méthode d'Euler améliorée avec $h = 0.1$ et tracer la solution numérique.

SOLUTION. Solution numérique par Matlab.— Voici la méthode d'Euler améliorée pour le problème en question :

```

clear
h = 0.1; x0= 0; xf= 1; y0 = 0;
n = ceil((xf-x0)/h); % nombre de pas
%
count = 2; print_time = 1; % quand ecrire les donnees de sortie
x = x0; y = y0; % initialiser x et y

```

```

output2 = [0 x0 y0];
for i=1:n
    zp = y + h*exr10_25(x,y); % methode d'Euler
    z = y + (1/2)*h*(exr10_25(x,y)+exr10_25(x+h,zp));
    x = x + h;
    if count > print_time
        output2 = [output2; i x z];
        count = count - print_time;
    end
    y = z;
count = count + 1;
end
output2
save output2 % pour imprimer le graphe

```

La commande output2 imprime les valeurs de n , x et y .

| n | x | y |
|-------------------|------------------|------------------|
| 0 | 0 | 0 |
| 1.00000000000000 | 0.10000000000000 | 0.10475020826390 |
| 2.00000000000000 | 0.20000000000000 | 0.21833345972227 |
| 3.00000000000000 | 0.30000000000000 | 0.33935117091202 |
| 4.00000000000000 | 0.40000000000000 | 0.46622105817179 |
| 5.00000000000000 | 0.50000000000000 | 0.59727677538612 |
| 6.00000000000000 | 0.60000000000000 | 0.73088021271199 |
| 7.00000000000000 | 0.70000000000000 | 0.86552867523997 |
| 8.00000000000000 | 0.80000000000000 | 0.99994084307400 |
| 9.00000000000000 | 0.90000000000000 | 1.13311147003613 |
| 10.00000000000000 | 1.00000000000000 | 1.26433264384505 |

□

Ex. 10.13. Soit le problème à valeur initiale

$$y' = x + \cos y, \quad y(0) = 0, \quad 0 \leq x \leq 1.$$

Approcher la solution à 6 décimales près par la méthode de Runge–Kutta d'ordre 4 avec $h = 0.1$ et tracer la solution numérique.

SOLUTION. Solution numérique par Matlab.—

```

clear
h = 0.1; x0= 0; xf= 1; y0 = 0;
n = ceil((xf-x0)/h); % nombre de pas
%
count = 2; print_time = 1; % quand ecrire les donnees de sortie
x = x0; y = y0; % initialiser x et y
output3 = [0 x0 y0];
for i=1:n
    k1 = h*exr10_25(x,y);
    k2 = h*exr10_25(x+h/2,y+k1/2);
    k3 = h*exr10_25(x+h/2,y+k2/2);
    k4 = h*exr10_25(x+h,y+k3);

```

```

z = y + (1/6)*(k1+2*k2+2*k3+k4);
x = x + h;
if count > print_time
    output3 = [output3; i x z];
    count = count - print_time;
end
y = z;
count = count + 1;
end
output3
save output3 % pour imprimer le graphe

```

La commande output3 imprime les valeurs de n , x et y .

| n | x | y |
|-------------------|------------------|------------------|
| 0 | 0 | 0 |
| 1.00000000000000 | 0.10000000000000 | 0.10482097362427 |
| 2.00000000000000 | 0.20000000000000 | 0.21847505355285 |
| 3.00000000000000 | 0.30000000000000 | 0.33956414151249 |
| 4.00000000000000 | 0.40000000000000 | 0.46650622608728 |
| 5.00000000000000 | 0.50000000000000 | 0.59763447559658 |
| 6.00000000000000 | 0.60000000000000 | 0.73130914485224 |
| 7.00000000000000 | 0.70000000000000 | 0.86602471267959 |
| 8.00000000000000 | 0.80000000000000 | 1.00049620051241 |
| 9.00000000000000 | 0.90000000000000 | 1.13371450064800 |
| 10.00000000000000 | 1.00000000000000 | 1.26496830711844 |

□

Ex. 10.25. Soit le problème à valeur initiale

$$y' = x + \cos y, \quad y(0) = 0, \quad 0 \leq x \leq 1.$$

Approcher la solution à 6 décimales près par le prédicteur-correcteur d'Adams–Bashforth–Moulton d'ordre 4 avec $h = 0.1$, estimer l'erreur locale en $x = 0.5$ et tracer la solution numérique.

SOLUTION. Solution numérique par Matlab.— On obtient les 4 valeurs de départ pour ABM4 au moyen des conditions initiales et de la méthode de Runge–Kutta d'ordre 4.

```

clear
h = 0.1; x0= 0; xf= 1; y0 = 0;
n = ceil((xf-x0)/h); % nombre de pas
%
count = 2; print_time = 1; % quand ecrire les donnees de sortie
x = x0; y = y0; % initialiser x et y
output4 = [0 x0 y0 0];
%RK4
for i=1:3
    k1 = h*exr10_25(x,y);
    k2 = h*exr10_25(x+h/2,y+k1/2);
    k3 = h*exr10_25(x+h/2,y+k2/2);

```

```

k4 = h*exr10_25(x+h,y+k3);
z = y + (1/6)*(k1+2*k2+2*k3+k4);
x = x + h;
    if count > print_time
        output4 = [output4; i x z 0];
        count = count - print_time;
    end
y = z;
count = count + 1;
end
% ABM4
for i=4:n
    zp = y + (h/24)*(55*exr10_25(output4(i,2),output4(i,3))-...
        59*exr10_25(output4(i-1,2),output4(i-1,3))+...
        37*exr10_25(output4(i-2,2),output4(i-2,3))-...
        9*exr10_25(output4(i-3,2),output4(i-3,3)) );
    z = y + (h/24)*( 9*exr10_25(x+h,zp)+...
        19*exr10_25(output4(i,2),output4(i,3))-...
        5*exr10_25(output4(i-1,2),output4(i-1,3))+...
        exr10_25(output4(i-2,2),output4(i-2,3)) );
    x = x + h;
    if count > print_time
        errest = -(19/270)*(z-zp);
        output4 = [output4; i x z errest];
        count = count - print_time;
    end
    y = z;
count = count + 1;
end
output4
save output4 % pour faire le graphe

```

La commande `output4` imprime les valeurs de n , x et y .

| n | x | y | Error estimate |
|--------------------|-------------------|------------------|-------------------|
| 0 | 0 | 0 | 0 |
| 1.000000000000000 | 0.100000000000000 | 0.10482097362427 | 0 |
| 2.000000000000000 | 0.200000000000000 | 0.21847505355285 | 0 |
| 3.000000000000000 | 0.300000000000000 | 0.33956414151249 | 0 |
| 4.000000000000000 | 0.400000000000000 | 0.46650952510670 | -0.00000234408483 |
| 5.000000000000000 | 0.500000000000000 | 0.59764142006542 | -0.00000292485029 |
| 6.000000000000000 | 0.600000000000000 | 0.73131943222018 | -0.00000304450366 |
| 7.000000000000000 | 0.700000000000000 | 0.86603741396612 | -0.00000269077058 |
| 8.000000000000000 | 0.800000000000000 | 1.00050998975914 | -0.00000195879670 |
| 9.000000000000000 | 0.900000000000000 | 1.13372798977088 | -0.00000104794662 |
| 10.000000000000000 | 1.000000000000000 | 1.26498035231682 | -0.00000017019624 |

□

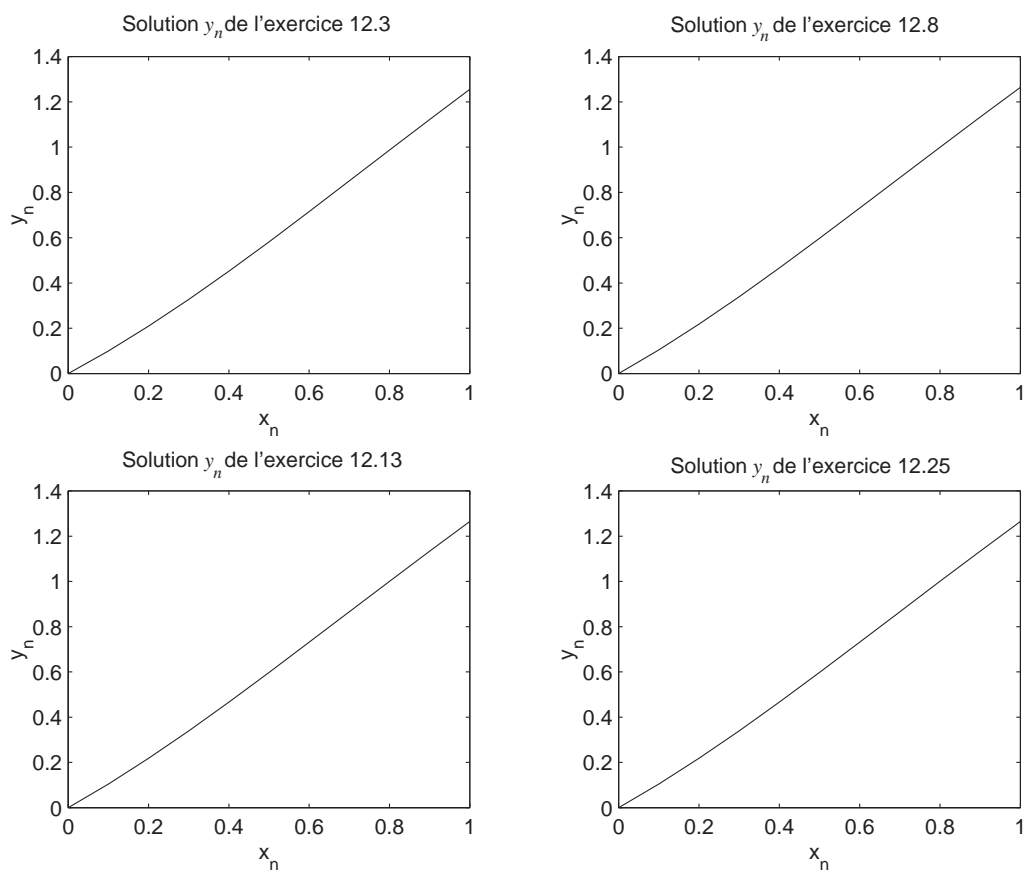


FIGURE 14.2. Solutions y_n des exercices 10.3 (Euler), 10.8 (Euler amélioré), 10.13 (RK4) et 10.25 (ABM4).

Les commandes suivantes tracent les solutions numériques des exercices 10.3, 10.8, 10.13 et 10.25.

```
load output1; load output2; load output3; load output4;
subplot(2,2,1); plot(output1(:,2),output1(:,3));
title('Solution y_n de l'exercice 10.3');
xlabel('x_n'); ylabel('y_n');
subplot(2,2,2); plot(output2(:,2),output2(:,3));
title('Solution y_n de l'exercice 10.8');
xlabel('x_n'); ylabel('y_n');
subplot(2,2,3); plot(output3(:,2),output3(:,3));
title('Solution y_n de l'exercice 10.13');
xlabel('x_n'); ylabel('y_n');
subplot(2,2,4); plot(output4(:,2),output4(:,3));
title('Solution y_n de l'exercice 10.25');
xlabel('x_n'); ylabel('y_n');
print -deps Fig9_3
```

Solutions des exercices du chapitre 11

Ex. 11.2. Solve the linear system

$$\begin{aligned}x_1 + x_2 + x_3 &= 5 \\x_1 + 2x_2 + 2x_3 &= 6 \\x_1 + 2x_2 + 3x_3 &= 8\end{aligned}$$

by the LU decomposition **without pivoting**.

SOLUTION. Solution numérique par Matlab.— In this case, Matlab need not pivot since L will be unit lower triangular. Hence we can use the LU decomposition obtained by Matlab.

```
clear
>>A = [1 1 1; 1 2 2; 1 2 3]; b = [5 6 8]';
>> [L,U] = lu(A) % LU decomposition of A
L =
     1     0     0
     1     1     0
     1     1     1
U =
     1     1     1
     0     1     1
     0     0     1
>> y = L\b % solution by forward substitution
y =
     5
     1
     2
>> x = U\y % solution by substitution retrograde
x =
     4
    -1
     2
```

□

Ex. 11.4. Solve the linear system

$$\begin{aligned}3x_1 + 9x_2 + 6x_3 &= 23 \\18x_1 + 48x_2 - 39x_3 &= 136 \\9x_1 - 27x_2 + 42x_3 &= 45\end{aligned}$$

by the LU decomposition **with pivoting**.

SOLUTION. Solution numérique par Matlab.— In this case, Matlab will pivot since L will be a row permutation of a unit lower triangular matrix. Hence we can use the LU decomposition obtained by Matlab.

```
clear
>>A = [3 9 6; 18 48 -39; 9 -27 42]; b = [23 136 45]';
>> [L,U] = lu(A) % LU decomposition of A
L =
     0.1667    -0.0196     1.0000
```

```

      1.0000      0      0
      0.5000      1.0000      0
U =
      18.0000      48.0000     -39.0000
           0     -51.0000      61.5000
           0           0      13.7059
>> y = L\b % solution by forward substitution
y =
      136.0000
      -23.0000
       -0.1176
>> x = U\y % solution by backward substitution
x =
       6.3619
       0.4406
      -0.0086

```

□

Ex. 11.6. Find the inverse of the Gaussian transformation

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -a & 1 & 0 & 0 \\ -b & 0 & 1 & 0 \\ -c & 0 & 0 & 1 \end{bmatrix}.$$

SOLUTION. The inverse, M^{-1} , of a Gaussian transformation is obtained by changing the signs of the multipliers, that is, of $-a, -b, -c$. Thus

$$M^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & 0 & 1 & 0 \\ c & 0 & 0 & 1 \end{bmatrix}. \quad \square$$

Ex. 11.7. Find the product of the three Gaussian transformations

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & 0 & 1 & 0 \\ c & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & d & 1 & 0 \\ 0 & e & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & f & 1 \end{bmatrix}.$$

SOLUTION. The product of three Gaussian transformations, $M_1M_2M_3$, in the given order is the unit lower triangular matrix whose j th column is the j th column of M_j .

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & d & 1 & 0 \\ c & e & f & 1 \end{bmatrix}. \quad \square$$

Ex. 11.10. Solve the linear system

$$\begin{bmatrix} 4 & 10 & 8 \\ 10 & 26 & 26 \\ 8 & 26 & 61 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 44 \\ 128 \\ 214 \end{bmatrix}.$$

by the Cholesky decomposition.

SOLUTION. The Matlab command `chol` decomposes a positive definite matrix A in the form

$$A = R^T R, \quad \text{where } R \text{ is upper triangular.}$$

```
>> A = [4 10 8; 10 26 26; 8 26 61]; b = [44 128 214]';
>> R = chol(A) % Cholesky decomposition
R =
     2     5     4
     0     1     6
     0     0     3
>> y = R\b % forward substitution
y =
    22
    18
     6
>> x = R\y % backward substitution
x =
    -8
     6
     2
```

□

Ex. 11.11. Do three iterations of Gauss–Seidel’s scheme on the properly permuted system with given initial vector $\mathbf{x}^{(0)}$,

$$\begin{aligned} 6x_1 + x_2 - x_3 &= 3 & x_1^{(0)} &= 1, \\ -x_1 + x_2 + 7x_3 &= -17 & x_2^{(0)} &= 1, \\ x_1 + 5x_2 + x_3 &= 0 & x_3^{(0)} &= 1. \end{aligned} \quad \text{with}$$

SOLUTION. Interchanging rows 2 and 3 and solving for x_1 , x_2 and x_3 , we have

$$\begin{aligned} x_1^{(n+1)} &= \frac{1}{6} [3 - x_2^{(n)} + x_3^{(n)}] & x_1^{(0)} &= 1, \\ x_2^{(n+1)} &= \frac{1}{5} [0 - x_1^{(n+1)} - x_3^{(n)}] & x_2^{(0)} &= 1, \\ x_3^{(n+1)} &= \frac{1}{7} [-17 + x_1^{(n+1)} - x_2^{(n+1)}] & x_3^{(0)} &= 1. \end{aligned} \quad \text{with}$$

Hence,

$$\mathbf{x}^{(1)} = \begin{bmatrix} 0.5 \\ -0.3 \\ -2.31429 \end{bmatrix}, \quad \mathbf{x}^{(2)} = \begin{bmatrix} 0.16428 \\ 0.43000 \\ -2.46653 \end{bmatrix}, \quad \mathbf{x}^{(3)} = \begin{bmatrix} 0.01724 \\ 0.48986 \\ -2.49609 \end{bmatrix}.$$

One suspects that

$$\mathbf{x}^{(n)} \rightarrow \begin{bmatrix} 0.0 \\ 0.5 \\ -2.5 \end{bmatrix}$$

as $n \rightarrow \infty$.

□

Ex. 11.14. Using least squares, fit a parabola to the data

$$(-1, 2), \quad (0, 0), \quad (1, 1), \quad (2, 2).$$

SOLUTION. We look for a solution of the form

$$f(x) = a_0 + a_1x + a_2x^2.$$

```
>> x = [-1 0 1 2]';
>> A = [x.^0 x x.^2];
>> y = [2 0 1 2]';
>> a = (A'*A \ (A'*y))'
a =
    0.4500    -0.6500    0.7500
```

The parabola is

$$f(x) = 0.45 - 0.65x + 0.75x^2.$$

The Matlab command `A \ y` produces the same answer. It uses the normal equations with the Cholesky or LU decomposition, or, perhaps, the QR decomposition, \square

Ex. 11.18. Determine and sketch the Gershgorin disks that contain the eigenvalues of the matrix

$$A = \begin{bmatrix} -2 & 1/2 & i/2 \\ 1/2 & 0 & i/2 \\ -i/2 & -i/2 & 2 \end{bmatrix}.$$

SOLUTION. The centres, c_i , and radii, r_i , of the disks are

$$\begin{aligned} c_1 &= -2, & r_1 &= |1/2| + |i/2| = 1, \\ c_2 &= 0, & r_2 &= |1/2| + |i/2| = 1, \\ c_3 &= 2, & r_3 &= |-i/2| + |i/2| = 1. \end{aligned}$$

Note that the eigenvalues are real since the matrix A is symmetric, $A^T = A$. \square

Formulaire et tables

15.1. Facteur d'intégration de $M(x, y) dx + N(x, y) dy = 0$

Soit l'équation différentielle homogène du 1er ordre

$$M(x, y) dx + N(x, y) dy = 0. \quad (15.1)$$

Si

$$\frac{1}{N} \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} \right) = f(x)$$

est une fonction de x seulement, alors

$$\mu(x) = e^{\int f(x) dx}$$

est un facteur d'intégration de (15.1).

Si

$$\frac{1}{M} \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} \right) = g(y)$$

est une fonction de y seulement, alors

$$\mu(y) = e^{-\int g(y) dy}$$

est un facteur d'intégration de (15.1).

15.2. Les polynômes de Legendre $P_n(x)$ sur $[-1, 1]$

1. L'équation différentielle de LEGENDRE :

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0, \quad -1 \leq x \leq 1.$$

2. La solution $y(x) = P_n(x)$ explicite :

$$P_n(x) = \frac{1}{2^n} \sum_{m=0}^{[n/2]} (-1)^m \binom{n}{m} \binom{2n-2m}{n} x^{n-2m},$$

où $[n/2]$ désigne le plus grand entier au plus égal à $n/2$.

3. La relation de récurrence :

$$(n + 1)P_{n+1}(x) = (2n + 1)xP_n(x) - nP_{n-1}(x).$$

4. La standardisation :

$$P_n(1) = 1.$$

5. La norme au carré des $P_n(x)$:

$$\int_{-1}^1 [P_n(x)]^2 dx = \frac{2}{2n + 1}.$$

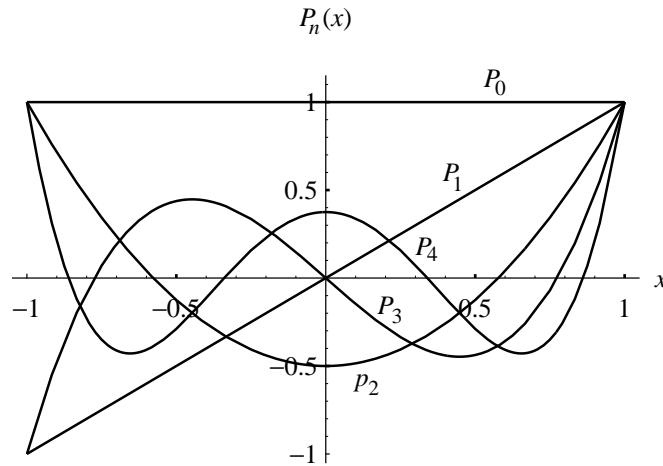


FIGURE 15.1. Les 5 premiers polynômes de LEGENDRE.

6. La formule de RODRIGUES :

$$P_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} \left[(1-x^2)^n \right].$$

7. La fonction génératrice :

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n, \quad -1 < x < 1, |t| < 1.$$

8. L'inégalité sur les $P_n(x)$:

$$|P_n(x)| \leq 1, \quad -1 \leq x \leq 1.$$

9. Les 6 premiers polynômes de LEGENDRE (V. figure 15.1) :

$$\begin{aligned} P_0(x) &= 1, & P_1(x) &= x, \\ P_2(x) &= \frac{1}{2}(3x^2 - 1), & P_3(x) &= \frac{1}{2}(5x^3 - 3x), \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3), & P_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x). \end{aligned}$$

15.3. Les polynômes de Laguerre sur $0 \leq x < \infty$

On définit les polynômes de LAGUERRE sur $0 \leq x < \infty$ par l'expression :

$$L_n(x) = \frac{e^x}{n!} \frac{d^n (x^n e^{-x})}{dx^n}, \quad n = 0, 1, \dots$$

Les 4 premiers sont (V. figure 15.2) :

$$\begin{aligned} L_0(x) &= 1, & L_1(x) &= 1 - x, \\ L_2(x) &= 1 - 2x + \frac{1}{2}x^2, & L_3(x) &= 1 - 3x + \frac{3}{2}x^2 - \frac{1}{6}x^3. \end{aligned}$$

On peut obtenir les $L_n(x)$ par la récurrence :

$$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x).$$

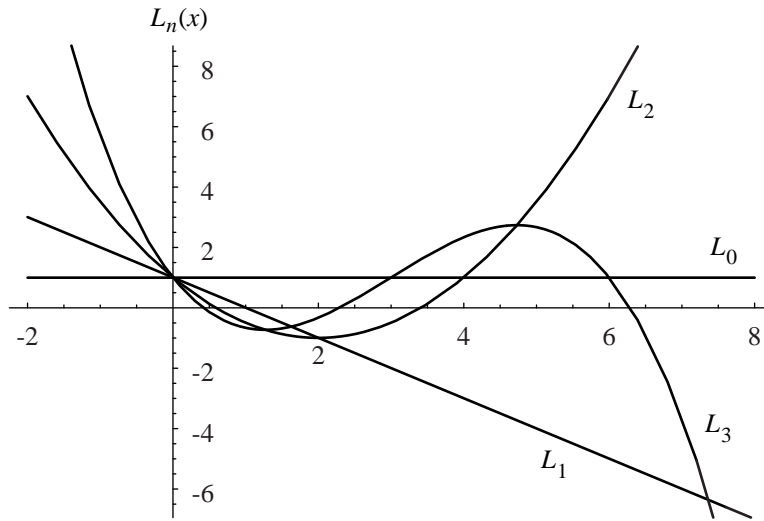


FIGURE 15.2. Les 4 premiers polynômes de LAGUERRE.

Les $L_n(x)$ sont solutions de l'équation différentielle

$$xy'' + (1-x)y' + ny = 0$$

et satisfont les relations d'orthogonalité avec le poids $p(x) = e^{-x}$:

$$\int_0^{\infty} e^{-x} L_m(x) L_n(x) dx = \begin{cases} 0, & m \neq n, \\ 1, & m = n. \end{cases}$$

15.4. Développements de Fourier-Legendre

Le développement de FOURIER-LEGENDRE de $f(\varphi)$ suivant les $P_n(\cos \varphi)$:

$$f(\varphi) \sim \sum_{n=0}^{\infty} a_n P_n(\cos \varphi), \quad 0 \leq \varphi \leq \pi,$$

où

$$\begin{aligned} a_n &= \frac{2n+1}{2} \int_0^{\pi} f(\varphi) P_n(\cos \varphi) \sin \varphi d\varphi \\ &= \frac{2n+1}{2} \int_{-1}^1 f(\arccos w) P_n(w) dw, \quad n = 0, 1, 2, \dots \end{aligned}$$

Le développement de FOURIER-LEGENDRE découle des relations d'orthogonalité suivantes :

$$\int_{-1}^1 P_m(x) P_n(x) dx = \begin{cases} 0, & m \neq n, \\ \frac{2}{2n+1}, & m = n. \end{cases}$$

TABLE 15.1. Table d'intégrales.

| | |
|-----|---|
| 1. | $\int \tan u \, du = \ln \sec u + c$ |
| 2. | $\int \cot u \, du = \ln \sin u + c$ |
| 3. | $\int \sec u \, du = \ln \sec u + \tan u + c$ |
| 4. | $\int \csc u \, du = \ln \csc u - \cot u + c$ |
| 5. | $\int \tanh u \, du = \ln \cosh u + c$ |
| 6. | $\int \coth u \, du = \ln \sinh u + c$ |
| 7. | $\int \frac{du}{\sqrt{a^2 - u^2}} = \arcsin \frac{u}{a} + c$ |
| 8. | $\int \frac{du}{\sqrt{a^2 + u^2}} = \ln \left(u + \sqrt{u^2 + a^2} \right) + c = \operatorname{arcsinh} \frac{u}{a} + c$ |
| 9. | $\int \frac{du}{\sqrt{a^2 - u^2}} = \ln \left(u + \sqrt{u^2 - a^2} \right) + c = \operatorname{arccosh} \frac{u}{a} + c$ |
| 10. | $\int \frac{du}{a^2 + u^2} = \frac{1}{a} \arctan \frac{u}{a} + c$ |
| 11. | $\int \frac{du}{u^2 - a^2} = \frac{1}{2a} \ln \left \frac{u - a}{u + a} \right + c$ |
| 12. | $\int \frac{du}{a^2 - u^2} = \frac{1}{2a} \ln \left \frac{u + a}{u - a} \right + c$ |
| 13. | $\int \frac{du}{u(a + bu)} = \frac{1}{a} \ln \left \frac{u}{a + bu} \right + c$ |
| 14. | $\int \frac{du}{u^2(a + bu)} = -\frac{1}{au} + \frac{b}{a^2} \ln \left \frac{a + bu}{u} \right + c$ |
| 15. | $\int \frac{du}{u(a + bu)^2} = \frac{1}{a(a + bu)} - \frac{1}{a^2} \ln \left \frac{a + bu}{u} \right + c$ |
| 16. | $\int x^n \ln ax \, dx = \frac{x^{n+1}}{n+1} \ln ax - \frac{x^{n+1}}{(n+1)^2} + c$ |

15.5. Table d'intégrales**15.6. La transformée de Laplace**

$$\mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} f(t) \, dt = F(s)$$

TABLE 15.2. Table de transformées de Laplace.

| | $F(s) = \mathcal{L}\{f(t)\}$ | $f(t)$ |
|-----|--|--|
| 1. | $F(s - a)$ | $e^{at} f(t)$ |
| 2. | $F(as + b)$ | $\frac{1}{a} e^{-bt/a} f\left(\frac{t}{a}\right)$ |
| 3. | $\frac{1}{s} e^{-cs}, c > 0$ | $h(t - c) := \begin{cases} 0, & 0 \leq t < c \\ 1, & t \geq c \end{cases}$ |
| 4. | $e^{-cs} F(s), c > 0$ | $f(t - c)h(t - c)$ |
| 5. | $F_1(s)F_2(s)$ | $\int_0^t f_1(\tau)f_2(t - \tau) d\tau$ |
| 6. | $\frac{1}{s}$ | 1 |
| 7. | $\frac{1}{s^{n+1}}$ | $\frac{t^n}{n!}$ |
| 8. | $\frac{1}{s^{a+1}}$ | $\frac{t^a}{\Gamma(a + 1)}$ |
| 9. | $\frac{1}{\sqrt{s}}$ | $\frac{1}{\sqrt{\pi t}}$ |
| 10. | $\frac{1}{s + a}$ | e^{-at} |
| 11. | $\frac{1}{(s + a)^{n+1}}$ | $\frac{t^n e^{-at}}{n!}$ |
| 12. | $\frac{k}{s^2 + k^2}$ | $\sin kt$ |
| 13. | $\frac{s}{s^2 + k^2}$ | $\cos kt$ |
| 14. | $\frac{k}{s^2 - k^2}$ | $\sinh kt$ |
| 15. | $\frac{s}{s^2 - k^2}$ | $\cosh kt$ |
| 16. | $\frac{2k^3}{(s^2 + k^2)^2}$ | $\sin kt - kt \cos kt$ |
| 17. | $\frac{2ks}{(s^2 + k^2)^2}$ | $t \sin kt$ |
| 18. | $\frac{1}{1 - e^{-ps}} \int_0^p e^{-st} f(t) dt$ | $f(t + p) = f(t), \text{ pour tout } t$ |

Index

- abscisse de convergence, 97
- absolument stable
 - méthode pour équadif, 223
- amplitude, 38

- base de Lagrange, 175
- BDF (backward differentiation formula), 245

- conditionnement d'une matrice, 266
- convergence d'une série
 - absolue, 125
 - uniforme, 125
- correcteur, 233
- critère
 - de convergence de D'ALEMBERT, 127
 - d'arrêt, 158

- décomposition
 - de Cholesky, 261
 - de Schur, 278
 - selon les valeurs singulières, 277
- différence divisée
 - k ème, 179
 - 1ère, 177
- différence prograde
 - k ème, 180
 - 1ère, 180
 - 2ème, 180
- différentiation rétrograde, 245
- Dormand–Prince
 - paire à sept étages, 227
- DP(5,4)7M, 227

- équation
 - de BERNOULLI, 21
 - de LEGENDRE, 130
- équation normale, 270
- erreur, 147
 - absolue, 147
 - d'arrondi, 147, 190, 212
 - de chute, 147
 - de méthode, 147, 210–212, 221
 - locale d'une méthode, 222
 - relative, 147
- étalonner les lignes, 259
- existence de solution analytique, 129
- extrapolation
 - de Richardson, 193
 - locale, 227, 233

- fonction
 - analytique, 129
 - d'ordre p , 210
 - de stabilité, 223
 - génératrice des $P_n(x)$, 135
 - lipschitzienne, 209
 - rationnelle, 125
 - sign, 149

- formule
 - centrée pour $f''(x)$, 190
 - centrée pour $f'(x)$, 190
 - de RODRIGUES des $P_n(x)$, 134
 - de différentiation à 3 points, 190
 - de différentiation numérique, 245
 - de différentiation prograde à 2 points, 189
 - de différentiation rétrograde à 2 points, 189

- frontière
 - à pente aux limites fixée, 186
 - libre, 186
 - naturelle, 186

- Gershgorin
 - disque de, 272
 - théorème de, 272

- intervalle de stabilité absolue, 222

- MATLAB
 - fonction `fzero`, 167
 - `ode113`, 243
 - `ode15i`, 251
 - `ode15s`, 251
 - `ode23`, 226
 - `ode23s`, 251
 - `ode23t`, 251
 - `ode23tb`, 251

- matrice
 - compagnon, 48
 - normale, 278
- matrice
 - à diagonale dominante, 262
 - définie positive, 261
- methode
 - méthode zéro-stable pour équadif, 222
- méthode
 - consistante pour équadif, 221
 - convergente pour équadif, 221
 - d'Adams–Bashforth–Moulton d'ordre 3, 234
 - d'Adams–Bashforth–Moulton d'ordre 4, 235
 - d'Euler, 210
 - d'Euler améliorée, 213
 - d'ordre p , 222
 - de Horner, 170
 - de Jacobi, 277
 - de la dichotomie, 151
 - de la parabole, 171
 - de la position fautive, 164
 - de la puissance, 274
 - inverse, 275
 - de la sécante, 163
 - de Lanczos, 277
 - de Müller, 171
 - de Newton, 159
 - de Newton modifiée, 162
 - de Newton–Raphson, 159
 - de Runge–Kutta d'ordre 3, 215
 - de Runge–Kutta d'ordre 4, 215, 216
 - de Runge–Kutta du 2ème ordre, 214
 - de Simpson, 196, 201
 - des points milieux, 198
 - des trapèzes, 199
 - du point milieu, 195
 - du trapèze, 195
 - FSAL, 228
 - globale Newton-bisection, 165
 - multipas, 232
 - multipas absolument stable, 232
 - multipas d'Adams–Bashforth, 232
 - multipas d'Adams–Moulton, 232
 - multipas explicite, 232, 242
 - multipas implicite, 232, 242
- mineur principal, 261
- NDF (numerical differentiation formula, 245
- norme
 - matricielle subordonnée, 266
- norme matricielle
 - l_1 , 266
 - l_2 , 266
 - l_∞ , 266
 - de Frobenius, 266
 - euclidienne, 266
- norme vectorielle
 - l_1 , 266
 - l_2 , 266
 - l_∞ , 266
 - supremum, 266
- numerical differentiation formula, 245
- opération
 - gaxpy, 265
 - saxpy, 264
- ordre de convergence
 - d'une récurrence, 159
- PECE
 - en mode, 234
- PECLE
 - en mode, 234
- période, 38
- phénomène de raideur, 244
- pivotage partiel, 253
- point fixe, 154
 - attractif, 154
 - indifférent, 154
 - répulsif, 154
- polynômes
 - de LAGUERRE $L_n(x)$, 326
 - de LEGENDRE $P_n(x)$, 132, 137, 325
- polynôme
 - d'interpolation
 - de Lagrange, 175
 - d'interpolation d'Hermite, 184
 - d'interpolation newtonnien aux différences divisées, 177
 - d'interpolation prograde de Gregory–Newton, 181
 - polynôme d'interpolation rétrograde de Gregory–Newton, 183
- prédicteur, 233
- problème
 - à valeur initiale du 1er ordre, 209
 - bien posé, 209
 - local, 222
- processus d'Aitken, 167
- QR
 - algorithme, 277
 - décomposition, 276
- quadrature gaussienne, 139, 206
 - à 2 points, 139, 206
 - à 3 points, 140, 206
- rapport de raideur, 244
- rayon de convergence d'une série, 127
- récurrence
 - de Bonnet, 290
 - de Gauss–Seidel, 267, 268
 - de Jacobi, 269
- réflexion de Householder, 276

- région de stabilité absolue, 222
- régression
 - linéaire, 270
 - quadratique, 271
- regula falsi, 164
- relation d'orthogonalité
 - des $L_n(x)$, 115, 327
 - des $P_n(x)$, 133, 327
- résidu, 276
- RKF(4,5), 230
- RKV(5,6), 231
- Runge–Kutta–Fehlberg
 - paire à 6 étages, 230
- Runge–Kutta–Verner
 - paire 8 étages, 231
- série de FOURIER–LEGENDRE, 136
- solution
 - exacte, 209
 - par récurrence, 267
- spline
 - aux pentes aux limites fixées, 186
 - du 3ème degré, 186
 - naturel, 186
- substitution
 - prograde, 255
 - rétrograde, 255
- système
 - raide, 243
- système
 - linéaire surdéterminé, 269
 - surdéterminé, 277
- table de différences divisées, 179
- tableau de Butcher, 215
- taux de convergence, 159
- théorème
 - des accroissements finis, 150
 - des valeurs extrêmes, 150
 - des valeurs intermédiaires, 150
 - des valeurs moyennes pour intégrales, 150
 - des valeurs moyennes pour les sommes, 150
- transformation gaussienne, 254
 - inverse, 254
 - produit, 255
- valeur propre d'une matrice, 272
- vecteur propre, 272
- virgule flottante, 147

Integration

$$\int uv' dx = uv - \int u'v dx$$

$$\int x^n dx = \frac{x^{n+1}}{n+1} + c \quad (n \neq -1)$$

$$\int \frac{1}{x} dx = \ln |x| + c$$

$$\int e^{ax} dx = \frac{1}{a} e^{ax} + c$$

$$\int \sin x dx = -\cos x + c$$

$$\int \cos x dx = \sin x + c$$

$$\int \tan x dx = -\ln |\cos x| + c$$

$$\int \cot x dx = \ln |\sin x| + c$$

$$\int \sec x dx = \ln |\sec x + \tan x| + c$$

$$\int \csc x dx = \ln |\csc x - \cot x| + c$$

$$\int \frac{dx}{x^2 + a^2} = \frac{1}{a} \arctan \frac{x}{a} + c$$

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin \frac{x}{a} + c$$

$$\int \frac{dx}{\sqrt{x^2 + a^2}} = \sinh^{-1} \frac{x}{a} + c$$

$$\int \frac{dx}{\sqrt{x^2 - a^2}} = \cosh^{-1} \frac{x}{a} + c$$

$$\int \sin^2 x dx = \frac{1}{2}x - \frac{1}{4}\sin 2x + c$$

$$\int \cos^2 x dx = \frac{1}{2}x + \frac{1}{4}\sin 2x + c$$

$$\int \tan^2 x dx = \tan x - x + c$$

$$\int \cot^2 x dx = -\cot x - x + c$$

$$\int \ln x dx = x \ln x - x + c$$

$$\begin{aligned} \int e^{ax} \sin bx dx \\ = \frac{e^{ax}}{a^2 + b^2} (a \sin bx - b \cos bx) + c \end{aligned}$$

$$\begin{aligned} \int e^{ax} \cos bx dx \\ = \frac{e^{ax}}{a^2 + b^2} (a \cos bx + b \sin bx) + c \end{aligned}$$

Laplace Transform: General Formulas

| Formula | Name, Comments |
|--|--|
| $F(s) = \mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} f(t) dt$ $f(t) = \mathcal{L}^{-1}\{F(s)\}$ | Definition of Transform Inverse Transform |
| $\mathcal{L}\{af(t) + bg(t)\} = a\mathcal{L}\{f(t)\} + b\mathcal{L}\{g(t)\}$ | Linearity |
| $\mathcal{L}\{e^{at}f(t)\} = F(s - a)$ $\mathcal{L}^{-1}\{F(s - a)\} = e^{at}f(t)$ | s-Shifting (First Shifting Theorem) |
| $\mathcal{L}\{f'\} = s\mathcal{L}\{f\} - f(0)$ $\mathcal{L}\{f''\} = s^2\mathcal{L}\{f\} - sf(0) - f'(0)$ $\mathcal{L}\{f^{(n)}\} = s^n\mathcal{L}\{f\} - s^{n-1}f(0) - \dots - f^{(n-1)}(0)$ $\mathcal{L}\left\{\int_0^t f(\tau) d\tau\right\} = \frac{1}{s}\mathcal{L}\{f\}$ | Differentiation of Function Integration of Function |
| $\mathcal{L}\{f(t - a)u(t - a)\} = e^{-as}F(s)$ $\mathcal{L}^{-1}\{e^{-as}F(s)\} = f(t - a)u(t - a)$ | t-Shifting (Second Shifting Theorem) |
| $\mathcal{L}\{tf(t)\} = -F'(s)$ $\mathcal{L}\left\{\frac{f(t)}{t}\right\} = \int_s^{\infty} F(\bar{s}) d\bar{s}$ | Differentiation of Transform Integration of Transform |
| $(f * g)(t) = \int_0^t f(\tau)g(t - \tau) d\tau$ $= \int_0^t f(t - \tau)g(\tau) d\tau$ $\mathcal{L}\{f * g\} = \mathcal{L}\{f\}\mathcal{L}\{g\}$ | Convolution |
| $\mathcal{L}\{f\} = \frac{1}{1 - e^{-ps}} \int_0^p e^{-st} f(t) dt$ | f Periodic with Period p |

$$\sin x \sin y = \frac{1}{2}[-\cos(x + y) + \cos(x - y)]$$

$$\cos x \cos y = \frac{1}{2}[\cos(x + y) + \cos(x - y)]$$

$$\sin x \cos y = \frac{1}{2}[\sin(x + y) + \sin(x - y)]$$

| $F(s) = \mathcal{L}\{f(t)\}$ | $f(t)$ | $F(s) = \mathcal{L}\{f(t)\}$ | $f(t)$ |
|--|--|---|--|
| $1/s$ $1/s^2$ $1/s^n \quad (n = 1, 2, \dots)$ $1/\sqrt{s}$ $1/s^{3/2}$ $1/s^a \quad (a > 0)$ | 1 t $t^{n-1}/(n-1)!$ $1/\sqrt{\pi t}$ $2\sqrt{t/\pi}$ $t^{a-1}/\Gamma(a)$ | $\frac{s}{(s^2 + \omega^2)^2}$ $\frac{s^2}{(s^2 + \omega^2)^2}$ $\frac{s}{(s^2 + a^2)(s^2 + b^2)} \quad (a^2 \neq b^2)$ | $\frac{t}{2\omega} \sin \omega t$ $\frac{1}{2\omega} (\sin \omega t + \omega t \cos \omega t)$ $\frac{1}{b^2 - a^2} (\cos at - \cos bt)$ |
| $\frac{1}{s-a}$ $\frac{1}{(s-a)^2}$ $\frac{1}{(s-a)^n} \quad (n = 1, 2, \dots)$ $\frac{1}{(s-a)^k} \quad (k > 0)$ | e^{at} te^{at} $\frac{1}{(n-1)!} t^{n-1} e^{at}$ $\frac{1}{\Gamma(k)} t^{k-1} e^{at}$ | $\frac{1}{s^4 + 4k^4}$ $\frac{s}{s^4 + 4k^4}$ $\frac{1}{s^4 - k^4}$ $\frac{s}{s^4 - k^4}$ | $\frac{1}{4k^3} (\sin kt \cos kt - \cos kt \sinh kt)$ $\frac{1}{2k^2} \sin kt \sinh kt$ $\frac{1}{2k^3} (\sinh kt - \sin kt)$ $\frac{1}{2k^2} (\cosh kt - \cos kt)$ |
| $\frac{1}{(s-a)(s-b)} \quad (a \neq b)$ $\frac{s}{(s-a)(s-b)} \quad (a \neq b)$ | $\frac{1}{(a-b)} (e^{at} - e^{bt})$ $\frac{1}{(a-b)} (ae^{at} - be^{bt})$ | $\sqrt{s-a} - \sqrt{s-b}$ $\frac{1}{\sqrt{s+a}\sqrt{s+b}}$ $\frac{1}{\sqrt{s^2+a^2}}$ | $\frac{1}{2\sqrt{\pi t^3}} (e^{bt} - e^{at})$ $e^{-(a+b)t/2} I_0\left(\frac{a-b}{2}t\right)$ $J_0(at)$ |
| $\frac{1}{s^2 + \omega^2}$ $\frac{s}{s^2 + \omega^2}$ $\frac{1}{s^2 - a^2}$ $\frac{s}{s^2 - a^2}$ $\frac{1}{(s-a)^2 + \omega^2}$ $\frac{s-a}{(s-a)^2 + \omega^2}$ | $\frac{1}{\omega} \sin \omega t$ $\cos \omega t$ $\frac{1}{a} \sinh at$ $\cosh at$ $\frac{1}{\omega} e^{at} \sin \omega t$ $e^{at} \cos \omega t$ | $\frac{s}{(s-a)^{3/2}}$ $\frac{1}{(s^2 - a^2)^k} \quad (k > 0)$ e^{-as}/s e^{-as} | $\frac{1}{\sqrt{\pi t}} e^{at}(1 + 2at)$ $\frac{\sqrt{\pi}}{\Gamma(k)} \left(\frac{t}{2a}\right)^{k-1/2} I_{k-1/2}(at)$ $u(t-a)$ $\delta(t-a)$ |
| $\frac{1}{s(s^2 + \omega^2)}$ $\frac{1}{s^2(s^2 + \omega^2)}$ $\frac{1}{(s^2 + \omega^2)^2}$ | $\frac{1}{\omega^2} (1 - \cos \omega t)$ $\frac{1}{\omega^3} (\omega t - \sin \omega t)$ $\frac{1}{2\omega^3} (\sin \omega t - \omega t \cos \omega t)$ | $\frac{1}{\sqrt{s}} e^{-k/s}$ $\frac{1}{s^{3/2}} e^{k/s}$ $e^{-k\sqrt{s}} \quad (k > 0)$ | $J_0(2\sqrt{kt})$ $\frac{1}{\sqrt{\pi t}} \cos 2\sqrt{kt}$ $\frac{1}{\sqrt{\pi k}} \sinh 2\sqrt{kt}$ $\frac{k}{2\sqrt{\pi t^3}} e^{-k^2/t}$ |
| | | $\frac{1}{s} \ln s$ $\ln \frac{s-a}{s-b}$ | $-\ln t - \gamma \quad (\gamma \approx 0.5772)$ $\frac{1}{t} (e^{bt} - e^{at})$ |

INTEGRATING FACTORS

$$\frac{1}{N} \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} \right) = f(x) \quad \rightarrow \quad \exp \left(\int f(x) dx \right)$$

$$\frac{1}{M} \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} \right) = g(y) \quad \rightarrow \quad \exp \left(- \int g(y) dy \right)$$

FOR $M(x, y) dx + N(x, y) dy = 0$.