

Kit de survie de l'option B

① Méthode de Newton Filbet

Th (convergence locale dans $\mathbb{R}$) Soit $f \in \mathcal{C}^2([a, b], \mathbb{R})$ et $\bar{x} \in]a, b[$ tel que $f(\bar{x}) = 0$ et $f'(\bar{x}) \neq 0$. Alors il existe $\delta > 0$ tel que pour tout $x_0 \in [\bar{x} \pm \delta]$, la suite $(x_n)_{n \in \mathbb{N}}$ définie par récurrence à partir de x_0 par : $\forall n \geq 0, x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ soit bien définie et converge vers $\bar{x}$, avec $|x_{n+1} - \bar{x}| \leq C |x_n - \bar{x}|^2$.

Preuve :

① Par continuité de f' , il existe $\alpha > 0$ tq $\forall x \in [\bar{x} \pm \alpha], f'(x) \neq 0$

→ Intervalle sur lequel la f° d'itération a un sens.

② On note $F(x) = x - \frac{f(x)}{f'(x)}$. On cherche $\delta \in]0, \alpha[$ tel que si $x \in [\bar{x} \pm \delta]$, alors $F(x) \in [\bar{x} \pm \delta]$. On écrit :

$$F(x) - \bar{x} = \frac{(x - \bar{x})f'(x) - f(x) + f(\bar{x})}{f'(x)} = \frac{f''(y)(x - \bar{x})^2}{2f'(x)} \text{ avec } y \in [\bar{x}, x].$$

$$\text{Puis } |F(x) - \bar{x}| \leq \underbrace{\frac{\max |f''|}{2 \min |f'|}}_{C := \text{Lagrange}} |x - \bar{x}|^2 \text{ où max, min sont pris sur } [\bar{x} \pm \alpha]$$
$$\leq C \delta^2 < \delta \text{ si } \delta < \frac{1}{C}.$$

③ On fixe $\delta < \min(\alpha, \frac{1}{C})$. Soit $x_0 \in [\bar{x} \pm \delta]$. Par ce qui précède, la suite $(x_n)_n$ est bien définie car reste dans $[\bar{x} \pm \delta] \subset [\bar{x} \pm \alpha]$ où $f' \neq 0$, et on a $\forall n \in \mathbb{N}, |x_{n+1} - \bar{x}| \leq C |x_n - \bar{x}|^2$ ■

NB : • On peut avoir une cv globale ^(à droite) avec une hypothèse supplémentaire de convexité.

• Si $f'(\bar{x}) = 0$ mais $f''(\bar{x}) \neq 0$, il y a cv mais non quadratique.

Th (convergence locale dans $\mathbb{R}^d$) Soit U un ouvert de $\mathbb{R}^d$ et $f \in \mathcal{C}^2(U, \mathbb{R}^d)$. Soit $\bar{x} \in U$ tel que $f(\bar{x}) = 0$ et $df(\bar{x})$ est inversible. Alors il existe $\delta > 0$ tel que pour tout $x_0 \in B(\bar{x}, \delta)$, la suite $(x_n)_{n \in \mathbb{N}}$ définie par récurrence à partir de x_0 par : $\forall n \in \mathbb{N}$, $x_{n+1} = x_n - df(x_n)^{-1} \cdot f(x_n)$ est bien définie, converge vers $\bar{x}$ et $\forall n \in \mathbb{N}$, $\|x_{n+1} - \bar{x}\| \leq C \|x_n - \bar{x}\|^2$.

Preuve. Même méthode que sur $\mathbb{R}$ avec des outils de calcul différentiel.

① Il existe $\alpha > 0$ tq pour $x \in B(\bar{x}, \alpha)$, $df(x) \in GL(\mathbb{R}^d)$ car f est $\mathcal{C}^2$ et $GL(\mathbb{R}^d)$ est ouvert. Par le théorème d'inversion locale, il existe même $\beta > 0$ tel que f réalise un $\mathcal{C}^1$ -difféomorphisme local de $B(\bar{x}, \beta)$ sur son image. Ainsi, df^{-1} est continue (en $\bar{x}$ et) sur $B(\bar{x}, \beta)$. En fixant $\gamma < \beta$, df^{-1} est continue sur le compact $f(B(\bar{x}, \gamma))$ donc bornée sur celui-ci. f' d'it.

② On cherche $\delta \in]0, \gamma[$ tq $\forall x \in B(\bar{x}, \delta)$, $F(x) \in B(\bar{x}, \delta)$.

On écrit :

$$\begin{aligned} F(x) - \bar{x} &= df(x)^{-1} [df(x) \cdot (x - \bar{x}) - f(x) + f(\bar{x})] \\ &\stackrel{T-R}{=} df(x)^{-1} \cdot \int_0^1 (1-t) d^2 f(\bar{x} + t(x - \bar{x})) \cdot (x - \bar{x}, x - \bar{x}) dt. \end{aligned}$$

$$\begin{aligned} \|F(x) - \bar{x}\| &\leq \underbrace{\|df(x)^{-1}\|_{\infty}}_{\text{sur } f(B(\bar{x}, \gamma))} \underbrace{\|d^2 f\|_{\infty}}_{\text{sur } B(\bar{x}, \gamma)} \cdot \|x - \bar{x}\|^2 \cdot \frac{1}{2} \\ &\leq C \|x - \bar{x}\|^2 \leq C \delta^2 < \delta \text{ si } \delta < \frac{1}{C}. \end{aligned}$$

③ Idem.

② Méthode du gradient à pas optimal. Bermir, Bermir

Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$ et $b \in \mathbb{R}^n \setminus \{0\}$. La méthode du gradient à pas optimal est une méthode de résolution approchée de l'équation $Ax = b$ par minimisation de l'application $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}$

$$x \mapsto \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle.$$

Th. • L'application Φ admet un unique minimum local, qui est global, atteint en $\bar{x} \in \mathbb{R}^n$ tq $A\bar{x} = b$.

• Soit $a \in \mathbb{R}^n \setminus \{0\}$ et $(x_k)_{k \in \mathbb{N}}$ la suite définie par récurrence à partir de $x_0 = a$ par

$$\begin{cases} v_k = \nabla \Phi(x_k) & \text{direction de descente} \\ \alpha_k = \frac{\langle v_k, v_k \rangle}{\langle v_k, Av_k \rangle} & \text{si } v_k \neq 0, 0 \text{ sinon} \\ x_{k+1} = x_k - \alpha_k v_k \end{cases}$$

La suite $(x_k)_{k \in \mathbb{N}}$ converge vers $\bar{x}$, et de plus :

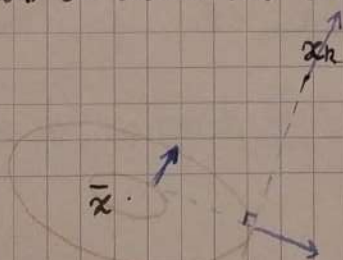
$$\|x_k - \bar{x}\| \leq \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right)^k \|x_0 - \bar{x}\|.$$

Preuve. • L'application Φ est de classe $\mathcal{C}^\infty$ et pour $x \in \mathbb{R}^n$:

$$\begin{cases} \nabla \Phi(x) = Ax - b \\ H\Phi(x) = A \in \mathcal{S}_n^{++}(\mathbb{R}) \end{cases}$$

Φ est strictement convexe et admet $\bar{x}$ comme unique point critique, donc Φ admet un unique minimum global atteint en $\bar{x}$, et c'est son unique extremum local.

• Le pas est choisi par minimisation de $f: t \mapsto \Phi(x_k - t \nabla \Phi(x_k))$. Les directions de descente sont orthogonales.



Lemme: $\forall x \in \mathbb{R}^n \setminus \{0\}$, $\frac{\|x\|^4}{\|x\|_A^2 \|x\|_{A^{-1}}^2} \geq 4 \frac{\lambda_{\max} \lambda_{\min}}{(\lambda_{\max} + \lambda_{\min})^2}$ orthonormée

Par le théorème spectral, on se place dans une base de vecteurs propre $(e_1, \dots, e_n)$ de A , associés aux v.p. $\lambda_1, \dots, \lambda_n$. Alors $(e_1, \dots, e_n)$ est aussi une base de vecteurs propres de A^{-1} associés aux v.p. $\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}$.

On écrit $x = \sum_{i=1}^n x_i e_i$ de sorte que :

$$\begin{aligned} \|x\|_A \|x\|_{A^{-1}} &= \left(\sum_{i=1}^n \lambda_i x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n \frac{1}{\lambda_i} x_i^2 \right)^{\frac{1}{2}} \\ &= \left(\frac{\lambda_{\max}}{\lambda_{\min}} \right)^{\frac{1}{2}} \left(\sum_{i=1}^n \frac{\lambda_i}{\lambda_{\max}} x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n \frac{\lambda_{\min}}{\lambda_i} x_i^2 \right)^{\frac{1}{2}} \\ &\stackrel{\text{Young}}{\leq} \frac{1}{2} \left(\frac{\lambda_{\max}}{\lambda_{\min}} \right)^{\frac{1}{2}} \sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_{\max}} + \frac{\lambda_{\min}}{\lambda_i} \right) x_i^2 \end{aligned}$$

L'application $g: [\lambda_{\min}, \lambda_{\max}] \rightarrow \mathbb{R}$ est convexe en

$$t \mapsto \frac{t}{\lambda_{\max}} + \frac{\lambda_{\min}}{t}$$

tant que somme de f° convexe. Or $g(\lambda_{\min}) = g(\lambda_{\max}) = 1 + \frac{\lambda_{\min}}{\lambda_{\max}}$ donc $\forall t \in [\lambda_{\min}, \lambda_{\max}]$, $g(t) \leq 1 + \frac{\lambda_{\min}}{\lambda_{\max}}$.

D'où :

$$\|x\|_A \|x\|_{A^{-1}} \leq \frac{1}{2} \left(\frac{\lambda_{\max}}{\lambda_{\min}} \right)^{\frac{1}{2}} \left(1 + \frac{\lambda_{\min}}{\lambda_{\max}} \right) \|x\|^2.$$

On obtient le lemme en élevant au carré.]

Contrôlons à présent $\|x_{k+1} - \bar{x}\|^2$, ou plutôt $\|x_{k+1} - \bar{x}\|_A^2$ qui est plus adaptée à notre étude. Comme on connaît $x_{k+1} - x_k$, c'est ce qu'on va faire apparaître. On suppose (x_k) non ultimat^o.

$$\|x_{k+1} - \bar{x}\|_A^2 = \langle A(x_{k+1} - \bar{x}), x_{k+1} - \bar{x} \rangle + \langle A(x_{k+1} - \bar{x}), x_k - \bar{x} \rangle$$

Or, on a vu que les directions de descente étaient ortho-

gonaler car $f'(x_k) = \langle \nabla \Phi(x_{k+1}), -\nabla \Phi(x_k) \rangle = 0$, donc :

$$\begin{aligned} \langle A(x_{k+1} - \bar{x}), x_{k+1} - x_k \rangle &= \langle Ax_{k+1} - b, -\alpha_k \nabla \Phi(x_k) \rangle \\ &= -\alpha_k \langle \nabla \Phi(x_{k+1}), \nabla \Phi(x_k) \rangle \\ &= 0. \end{aligned}$$

Puis, comme $A \in \mathcal{S}_n^+(\mathbb{R})$.

$$\begin{aligned} \|x_{k+1} - \bar{x}\|_A^2 &= \langle x_{k+1} - x_k, A(x_k - \bar{x}) \rangle + \underbrace{\langle x_k - \bar{x}, A(x_k - \bar{x}) \rangle}_{\|x_k - \bar{x}\|_A^2} \\ &= -\alpha_k \|\nabla \Phi(x_k)\|^2 + \|x_k - \bar{x}\|_A^2 \\ &= -\frac{\|v_k\|^4}{\|v_k\|_A^2} + \|x_k - \bar{x}\|_A^2 \end{aligned}$$

On remarque que $\|x_k - \bar{x}\|_A^2 = \langle A(x_k - \bar{x}), A^{-1} A(x_k - \bar{x}) \rangle$
 $= \|v_k\|_{A^{-1}}^2$.

d'où $\|x_{k+1} - \bar{x}\|_A^2 \leq \left(1 - \frac{\|v_k\|^4}{\|v_k\|_A^2 \|v_k\|_{A^{-1}}^2}\right) \|x_k - \bar{x}\|_A^2$.

Lemme

$$\begin{aligned} &\leq \left(1 - 4 \frac{\lambda_{\max} \lambda_{\min}}{(\lambda_{\min} + \lambda_{\max})^2}\right) \|x_k - \bar{x}\|_A^2 \\ &= \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^2 \|x_k - \bar{x}\|_A^2 \end{aligned}$$

La suite $(x_k)_{k \in \mathbb{N}}$ converge à vitesse géométrique vers $\bar{x}$.
 Pour se ramener à $\|\cdot\|$ à partir de $\|\cdot\|_A$ on écrit :

$$\sqrt{\lambda_{\min}} \|\cdot\| \leq \|\cdot\|_A \leq \sqrt{\lambda_{\max}} \|\cdot\|$$

NB: Interprétation de la constante.

Conditionnement. On a $\text{cond } A = \frac{\lambda_{\max}}{\lambda_{\min}}$ donc on peut écrire :

$$\|x_{k+1} - \bar{x}\| \leq \sqrt{\text{cond } A} \left(\frac{\text{cond } A - 1}{\text{cond } A + 1}\right)^{k+1} \|x_0 - \bar{x}\|^2$$

La convergence est plus rapide pour une matrice bien conditionnée.

Géométriquement

Si $n=2$, en notant λ_1 et λ_2 les vp de A , si on peut relier λ_1 et λ_2 à l'excentricité des ellipses des lignes de niveau. Faisons-le précisément. Soit $\alpha \in \mathbb{R}$ et $x \in \{\Phi(x) = \alpha\}$. On écrit $x = x_1 e_1 + x_2 e_2$ où (e_1, e_2) est une BON de $\mathbb{R}^2$ formée de v.p. de A . Alors, en écrivant $\bar{x} = \bar{x}_1 e_1 + \bar{x}_2 e_2$:

$$\Phi(x) = \alpha \Leftrightarrow \frac{1}{2} \langle Ax, x \rangle - \langle A\bar{x}, x \rangle = \alpha$$

$$\Leftrightarrow \lambda_1 x_1^2 + \lambda_2 x_2^2 - 2\lambda_1 \bar{x}_1 x_1 - 2\lambda_2 \bar{x}_2 x_2 = 2\alpha$$

$$\Leftrightarrow \lambda_1 (x_1 - \bar{x}_1)^2 + \lambda_2 (x_2 - \bar{x}_2)^2 - \underbrace{\lambda_1 \bar{x}_1^2 + \lambda_2 \bar{x}_2^2}_{\langle A\bar{x}, \bar{x} \rangle} = 2\alpha$$

$$\Leftrightarrow \left(\frac{x_1 - \bar{x}_1}{\sqrt{\lambda_1}} \right)^2 + \left(\frac{x_2 - \bar{x}_2}{\sqrt{\lambda_2}} \right)^2 = \frac{2\alpha + \langle A\bar{x}, \bar{x} \rangle}{\lambda_1 \lambda_2}$$

On obtient l'équation d'une ellipse de demi-axes $\sqrt{\lambda_1}$ et $\sqrt{\lambda_2}$.

Plus λ_1 et λ_2 sont proches, plus les lignes de niveau ressemblent à des cercles et plus la cv est rapide.



cv en une étape



④ Décomposition LU, de Cholevsky Fiket.

Th (Existence & unicité de la décomposition LU) Soit $A \in GL_n(K)$. On suppose que les mineurs fondamentaux de A sont non nuls. Alors il existe un unique couple $(L, U) \in M_n(K)^2$ tq $A = LU$, L est triangulaire inférieure avec des 1 sur la diagonale, U triangulaire supérieure.

Preuve: • Unicité: si $LU = \tilde{L}\tilde{U}$, $\tilde{L}^{-1}L = \tilde{U}U^{-1}$ est une matrice triangulaire inférieure avec des 1 sur la diagonale, et également triangulaire supérieure. Donc c'est l'identité.

• Existence: on montre par récurrence que les pivots de la méthode du pivot de Gauss appliquée à A ne sont jamais nuls. Le premier pivot est bien non nul puisque $a_{11} \neq 0$. Soit $k \in [1, n-1]$. On suppose que les k premières pivots étaient non nuls. Alors l'appliqué du pivot de Gauss fournit L_k triangulaire inférieure, telle que:

$$L_k A = \left[\begin{array}{c|c} \begin{array}{c} \diagup \\ (0) \end{array} & * \\ \hline (0) & \begin{array}{c} \text{pivot} \\ * \end{array} \end{array} \right]$$

avec des 1 sur la diagonale

Par calcul par bloc, $L_k^{(k+1)} \cdot A^{(k+1)} = \left[\begin{array}{c|c} \begin{array}{c} \diagup \\ (0) \end{array} & * \\ \hline 0 & \begin{array}{c} \text{pivot} \\ * \end{array} \end{array} \right]$

En prenant le déterminant, on a $\det(A^{(k+1)}) \times 1 = \text{pivot}_{k+1, k+1} \cdot \prod (\text{pivots})$
 D'où $\text{pivot}_{k+1, k+1} \neq 0$ ≠ 0 hyp. ≠ 0 HR

NB: Ce résultat s'applique en particulier aux matrices symétriques définies positives.

• Si $A \in GL_n(K)$, $\exists P$ permutat° tq PA admet une décompos° LU.

• On peut écrire $A = LDU$ avec D diagonale et L et U ayant des 1 sur leurs diagonales.

Th (Factorisation de Cholesky). Soit $A \in \mathcal{M}_n^+(\mathbb{R})$. Il existe une unique matrice R triangulaire supérieure à coefficients diagonaux > 0 tq $A = R^* R$.

Preuve: On écrit $A = LDU$. A étant symétrique définie positive, toutes ses sous-matrices principales sont symétriques définies positives et, en utilisant le déterminant, les coef de D sont strictement positifs. On écrit donc $D = \Delta \Delta$ avec $\Delta = \sqrt{D}$. On conclut en transposant: ${}^t A = {}^t U \Delta \Delta {}^t L$. Par unicité de la décomposition ${}^t U \Delta = L \Delta$ d'où $L = {}^t U$. Donc $R = \Delta U$ convient.

⑤ Méthodes itératives de résolution de systèmes linéaires. Filbet.

Soit $A \in GL_n(\mathbb{C})$ et $b \in \mathbb{C}^n$. On cherche à résoudre l'équation $Ax = b$ d'inconnue $x \in \mathbb{C}^n$. Plutôt que d'inverser directement A par pivot, on écrit $A = M - N$ avec $M \in GL_n(\mathbb{C})$ facile à inverser. Partant de $x_0 \in \mathbb{C}^n$, on définit la suite $(x_k)_{k \in \mathbb{N}}$ facile à calculer par $Mx_{k+1} = Nx_k + b$.

Pour de bons choix de M, N , cette suite peut converger vers la solution x .

On note $e_k = x_k - x$ l'erreur / à la solution. On a:

$$\begin{aligned} e_{k+1} &= x_{k+1} - x \\ &= M^{-1} N x_k + M^{-1} b - x \\ &= M^{-1} N x_k + M^{-1} (M - N) x - x \\ &= M^{-1} N (x_k - x) \\ &= M^{-1} N e_k. \end{aligned}$$

Ce qui nous importe est donc la convergence de $((M^{-1}N)^k)_{k \in \mathbb{N}}$, que l'on relie au rayon spectral $\rho(M^{-1}N)$.

Th: Soit $B \in \mathcal{M}_n(\mathbb{C})$.

(i) $B^k \xrightarrow[k \rightarrow \infty]{} 0$

(ii) Il existe $\|\cdot\|$ une norme subordonnée telle que $\|B\| < 1$.

(iii) $\rho(B) < 1$.

Preuve: Les implications (ii) $\Rightarrow$ (i) $\Rightarrow$ (iii) et (iii) $\Rightarrow$ (ii) sont claires. On va donc justifier (iii) $\Rightarrow$ (ii).

Supposons $\rho(B) < 1$. On trigonalise B : il existe $P \in GL_n(\mathbb{C})$ tq $P^{-1}BP = T = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ (0) & & \lambda_n \end{bmatrix}$ où $\lambda_1, \dots, \lambda_n$ sont v.p. de B .

Soit $\delta > 0$ à ajuster. On note $Q\delta = \begin{bmatrix} \delta & & \\ & \ddots & \\ & & \delta^{n-1} \end{bmatrix}$. Alors:

$$Q\delta^{-1}P^{-1}BPQ\delta = \begin{bmatrix} \lambda_1 \delta_{1,2} & & \delta^{n-1} t_{1,n} \\ & \ddots & \\ (0) & & \delta t_{n-1,n} \\ & & & \lambda_n \end{bmatrix} =: T\delta$$

On note $\|x\|_\delta = \|(PQ\delta)^{-1}x\|_1$ pour $x \in \mathbb{C}^n$.

Alors:

$$\begin{aligned} \|B\|_\delta &= \max_{\substack{x \in \mathbb{C}^n \setminus \{0\} \\ \|x\|_\delta = 1}} \|Bx\|_\delta = \max_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|(PQ\delta)^{-1}B(PQ\delta)(PQ\delta)^{-1}x\|_1}{\|(PQ\delta)^{-1}x\|_1} \\ &= \max_{y \in \mathbb{C}^n \setminus \{0\}} \frac{\|T\delta y\|_1}{\|y\|_1} \\ &= \|T\delta\|_1 \\ &= \max(|\lambda_1|, \dots, |\lambda_n| + \sum_{i=1}^{n-1} \delta |t_{i,n}|). \end{aligned}$$

On fixe δ de sorte que $\|T_\delta\|_1 < 1$.

On a donc établi que la méthode itérative converge vers la solution de $Ax=b$ pour tout vecteur initial $x_0 \in \mathbb{C}^m$ si et seulement si $\rho(M^{-1}N) < 1$.

Application à la méthode de Jacobi $M =$ diagonale de A .

Th: Soit $A \in M_n(\mathbb{C})$ à diagonale strictement dominante. Alors la méthode de Jacobi est convergente. $\hookrightarrow \forall i \in [1, n], |a_{ii}| > \sum_{j \neq i} |a_{ij}|$.

$$M^{-1}N = \begin{bmatrix} 0 & & -\frac{a_{1j}}{a_{11}} \\ & \ddots & \\ -\frac{a_{ij}}{a_{ii}} & & 0 \end{bmatrix}$$

$$\text{Donc } \|M^{-1}N\|_\infty = \max_{i \in [1, n]} \sum_{j=1}^m |(M^{-1}N)_{ij}| < 1.$$

Application à la méthode de Gauss-Seidel $M =$ partie tri. inf. de A .

Th: Soit $A \in \mathcal{P}_n^{++}(\mathbb{R})$. Alors la méthode de Gauss-Seidel est convergente.

On a besoin d'un lemme technique.

Lemme: Soit $A \in \mathcal{P}_n^{++}(\mathbb{R})$. On écrit $A = M - N$ avec $M \in GL_n(\mathbb{R})$.

Si $M+N \in \mathcal{P}_n^{++}(\mathbb{R})$, alors $\rho(M^{-1}N) < 1$.

Preuve: On montre que $\|M^{-1}N\|_A < 1$.

$$\|M^{-1}N\|_A^2 = \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\langle M^{-1}Nx, AM^{-1}Nx \rangle}{\langle x, Ax \rangle}$$

Soit $x \in \mathbb{R}^n \setminus \{0\}$. On écrit $N = -A + M$:

$$\langle M^{-1}Nx, AM^{-1}Nx \rangle$$

$$= \langle (I - M^{-1}A)x, A(I - M^{-1}A)x \rangle$$

$$= \langle x, Ax \rangle - 2\langle x, AM^{-1}Ax \rangle + \langle M^{-1}Ax, AM^{-1}Ax \rangle$$

> 0?

$$\text{Donc } \frac{\langle M^{-1}Nx, AM^{-1}Nx \rangle}{\langle x, Ax \rangle} = 1 - \frac{\overbrace{2\langle x, AM^{-1}Ax \rangle - \langle M^{-1}Ax, AM^{-1}Ax \rangle}}{\langle x, Ax \rangle}$$

Posons $y = M^{-1}Ax$. Alors :

$$2\langle x, AM^{-1}Ax \rangle - \langle M^{-1}Ax, AM^{-1}Ax \rangle$$

$$= 2\langle x, Ay \rangle - \langle y, Ay \rangle$$

$$= \langle 2Ax - Ay, y \rangle$$

$$= \langle 2My - Ay, y \rangle$$

$$= \langle (M + {}^tM)y - Ay, y \rangle \text{ par bilinéarité du produit scalaire.}$$

$$= \langle ({}^tM + N)y, y \rangle$$

> 0

Comme le sup est atteint, cela prouve $\|M^{-1}N\|_A < 1$. ▢

Preuve du th On vérifie d'abord que M la partie triangulaire supérieure de A est inversible, i.e. que la diagonale de A est constituée de coefficients strictement positifs. C'est bien le cas puisque pour $1 \leq i \leq n$, $a_{ii} = \langle e_i, Ae_i \rangle > 0$ car A est définie positive. diag. inf. sup

Vérifions que $M^T + N \in \mathcal{S}_n^{++}(\mathbb{R})$. On a, en notant $A = D + E + F$.

- $E^T = F$ car $A^T = A$

- $M = D + E$ donc $M^T = D + F$.

- $N = M - A = -F$

Donc $M^T + N = D$ est bien ^{symétrique} définie positive.

Par le lemme, $\rho(M^{-1}N) < 1$ et la méthode converge.

⑥ Equation de transport linéaire 1D Berthelin

On considère l'EDP linéaire d'ordre 1 :

$$(T) \begin{cases} \partial_t u(t, x) + a(t, x) \partial_x u(t, x) = 0. \\ u(0, \cdot) = u^0 \end{cases}$$

où $a \in \mathcal{C}^1(I \times \mathbb{R}, \mathbb{R})$, $u^0 \in \mathcal{C}^1(\mathbb{R} \times \mathbb{R})$, d'inconnue $u \in \mathcal{C}^1(I \times \mathbb{R}, \mathbb{R})$.

Analyse: La méthode des caractéristiques consiste à chercher des courbes le long desquelles u est constante.

Pour $x_0 \in \mathbb{R}$, soit $X(\cdot, x_0)$ la solution maximale de

$$\begin{cases} X'(t) = a(t, X(t)). \\ X(0) = x_0 \end{cases}$$

tg $X(\cdot, x_0)$ existe

Alors $\frac{d}{dt}(u(t, X(t))) = 0$, donc $\forall t, u(t, X(t, x_0)) = u^0(x_0)$.

Soit maintenant $(t, x) \in I \times \mathbb{R}$. Pour déterminer $u(t, x)$, on cherche $x_0 \in \mathbb{R}$ tel que $X(\cdot, x_0)$ soit définie en t et $X(t, x_0) = x$. Alors $u(t, x) = u(t, X(t, x_0)) = u^0(x_0)$.

Th. On suppose que $a \in \mathcal{C}^1(I \times \mathbb{R}, \mathbb{R})$ est globalement lipschitzienne selon sa deuxième variable ($\partial_x a \in L^\infty(I \times \mathbb{R}, \mathbb{R})$).

Alors (T) admet une unique solution $u \in \mathcal{C}^1(I \times \mathbb{R}, \mathbb{R})$, donnée par : $\forall (t, x) \in I \times \mathbb{R}$, $u(t, x) = u^0(X(0, t, x))$, où X est le flot de l'équation $X'(t) = a(t, X(t))$.

Preuve: ① La régularité sur a permet d'appliquer le théorème de Cauchy-Lipschitz global à l'EDO $X'(t) = a(t, X(t))$, donc le flot X est défini sur $I \times I \times \mathbb{R}$.

② On a $X(t, t_0, X(t_0, t, x)) = x$ pour tout $(t, t_0, x) \in I \times I \times \mathbb{R}$.

donc $V(t, t_0) \in I \times I$, $X(t, t_0, \cdot)$ est bijective d'inverse $X(t_0, t, \cdot)$.

③ L'unicité provient de: TFC le long des caractéristiques
 $V(t, x) \in I \times \mathbb{R}$, $u(t, x) = u(t, X(t, 0, X(0, t, x)))$
 $= u^0(X(0, t, x))$

④ Pour l'existence, on vérifie que la fonction définie dans l'unicité est solution. □

ex) Dans le cas $a(t, x) \equiv a$ constant, on a $X(t, t_0, x) = x + a(t - t_0)$, donc $u(t, x) = u^0(X(0, t, x)) = u^0(x - at)$.

Prop (régularité des solutions)

Hyp due th $\uparrow$

(i) Si $u^0, a \in \mathcal{C}^\infty$, alors $u \in \mathcal{C}^\infty$ [régularité du flot]

(ii) $u(t, \cdot)$ a la même régularité L^p que $u(t_0, \cdot)$ [chgt de var]
 $\Leftrightarrow$ en 1D

Détail pour (ii) Le cas $p = \infty$ est clair. On suppose $p < +\infty$.

$$\|u(t, \cdot)\|_p^p = \int_{\mathbb{R}} |u^0(X(0, t, x))|^p dx$$

$$\begin{matrix} y = X(0, t, x) \\ x = X(t, 0, y) \end{matrix} = \int_{\mathbb{R}} |u^0(y)|^p \left| \frac{\partial X}{\partial y}(t, 0, y) \right| dy$$

$$\begin{aligned} \text{Or } \partial_y X(t, 0, y) &= \lim_{h \rightarrow 0} \frac{1}{h} (X(t, 0, y+h) - X(t, 0, y)) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (y+h + \int_0^t a(s, y+h) ds - y - \int_0^t a(s, y) ds) \\ &= \lim_{h \rightarrow 0} \left(1 + \int_0^t \frac{a(s, y+h) - a(s, y)}{h} ds \right) \\ &\stackrel{\text{TCD}}{=} 1 + \int_0^t \partial_x a(s, y) ds \end{aligned}$$

$$\text{Donc } |\partial_y X(t, 0, y)| \leq 1 + \|\partial_x a\|_{L^\infty} t \leq \exp(t \| \partial_x a \|_{L^\infty}).$$

Finalement, $\|u(t, \cdot)\|_{L^p} \leq e^{\frac{1}{p}|t| \|\partial_x a\|_{L^\infty}} \|u^0\|_{L^p} < +\infty$

Pour résoudre le problème non homogène $\partial_t u + a \partial_x u = f$, regarder ce qui se passe le long des caractéristiques.

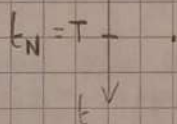
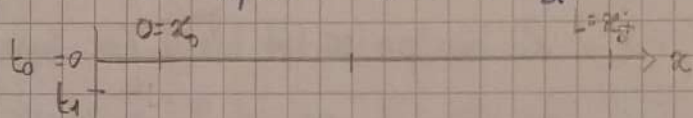
Schémas numériques:

On va étudier la convergence de deux schémas numériques sur l'équation de transport à coef constant:

$$\begin{cases} \partial_t u + a \partial_x u = 0 & \text{où } a > 0. \\ u(0, \cdot) = u^0 \end{cases}$$

On sait que $u(t, x) = u^0(x - at)$. [propagation vers la droite]

On fixe un domaine de résolution $[0, T] \times [0, L]$, un pas en temps Δt et en espace Δx , et $N = \frac{T}{\Delta t} \in \mathbb{N}^*$, $J = \frac{L}{\Delta x} \in \mathbb{N}^*$



On calcule à l'aide d'un schéma numérique $(u_j^n)_{j \in [0, J-1]}$ supposé approcher $(u(t_n, x_j))_{j \in [0, J-1]}$.

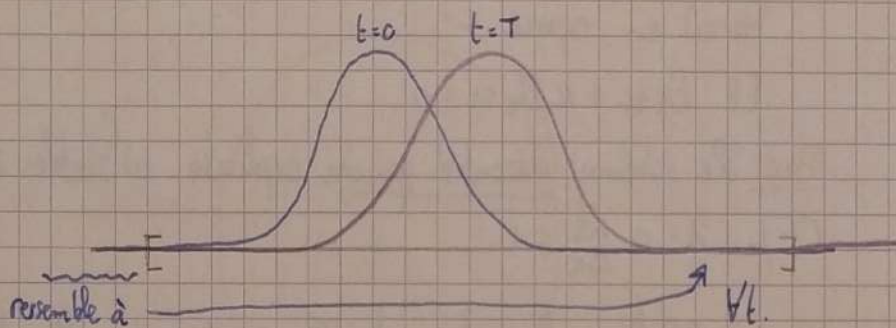
Schéma décentré à gauche explicite: $u_j^{n+1} = u_j^n - \Delta t a_0 \frac{u_j^n - u_{j-1}^n}{\Delta x}$

Schéma décentré à gauche implicite: $u_j^{n+1} = u_j^n - \Delta t a_0 \frac{u_j^{n+1} - u_{j-1}^{n+1}}{\Delta x}$

⚠ Il faut décentrer du côté d'où vient l'information, sinon le schéma sera instable.

On va mener une analyse L^2 à l'aide de la transformée de Fourier discrète. Pour cela, on doit imposer des conditions de bord périodiques: $u_0^n = u_J^n$, $u_{-1}^n = u_{J-1}^n$...

NB: Le choix peut paraître surprenant, surtout si la solution recherchée n'est pas vraiment périodique. Mais en pratique, on choisira un domaine en espace contenant «largement le support» de la solution pendant toute la durée de résolution.



Comme pour l'étude de schémas numériques d'EDO, on sépare l'analyse en une étape de consistance et une étape de stabilité.

Def: On dit que le schéma est consistant d'ordre k en temps et l en espace lorsque:

$$\sup_{m \in [0, N]} \sup_{j \in [0, J-1]} \underbrace{\left| \frac{u(t_{m+1}, x_j) - u(t_m, x_j)}{\Delta t} + a \frac{u(t_m, x_j) - u(t_m, x_{j-1})}{\Delta x} \right|}_{\varepsilon_j^m} = O(\Delta t^k + \Delta x^l).$$

Il peut y avoir des variat° sur la def.

Def: On dit que le schéma est L^2 -stable si $\forall (u_j^0)_{j \in [0, J-1]}$, il existe $C > 0$ ind. de Δt et Δx tq:

$$\forall N \in \mathbb{N}, \forall m \in [0, N], \underbrace{\frac{1}{\sqrt{J}} \|u^m\|_{\ell^2([0, J-1])}}_{\approx \|u(t_m, \cdot)\|_{L^2([0, L])}} \leq C \underbrace{\frac{1}{\sqrt{J}} \|u^0\|_{\ell^2([0, J-1])}}_{\approx \|u(t_0, \cdot)\|_{L^2([0, L])}}$$

Prop: Les deux schémas sont consistants d'ordre 1 en temps et 1 en espace.

Preuve: Taylor

$$\frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} = \partial_t u(t_n, x_j) + \int_0^1 \frac{1-s}{1!} \partial_t^2 u(t_n + s(t_{n+1} - t_n), x_j) (t_{n+1} - t_n) ds$$

$$\left| \frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} - \partial_t u(t_n, x_j) \right| \leq \frac{\Delta t}{2} \|\partial_t^2 u\|_{L^\infty([0, T] \times [0, L])}$$

Parcél en espace.

$$\text{Et } \partial_t u + a \partial_x u = 0.$$

Prop: Le schéma décentré gauche explicite est stable sous condition CFL $\lambda = a \frac{\Delta t}{\Delta x} \leq 1$.

Preuve: Utilisons l'analyse de Fourier discret. La transformée de Fourier discrète:

$$\mathcal{F}_J: \begin{cases} \mathbb{C}^J \rightarrow \mathbb{C}^J \\ (f_j)_j \mapsto \left(\frac{1}{\sqrt{J}} \sum_{j=0}^{J-1} f_j e^{-ik \frac{2\pi}{L} \frac{j}{J}} \right)_{k \in \left[-\lfloor \frac{J}{2} \rfloor, \lfloor \frac{J}{2} \rfloor - 1 \right]} \end{cases}$$

est une isomorphisme, et si on note $C^n = \mathcal{F}_J(U^n)$, on a:

$$\|C^n\|_{\ell^2} = \frac{1}{\sqrt{J}} \|U^n\|_{\ell^2}.$$

$$\text{Pour } C_k^{n+1} = \left(1 - a_0 \frac{\Delta t}{\Delta x} \left(1 - e^{-ik \frac{2\pi}{L} \Delta x} \right) \right) C_k^n$$

Le schéma est donc stable si:

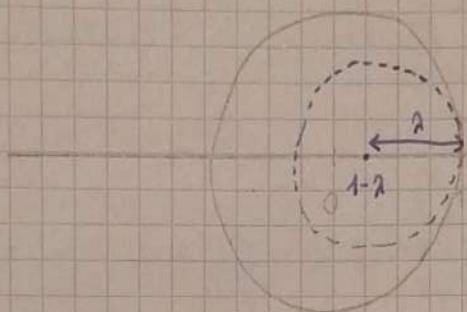
$$\forall k, \forall m \in [0, J-1], \underbrace{\left| 1 - a_0 \frac{\Delta t}{\Delta x} \left(1 - e^{-ik \frac{2\pi}{L} \Delta x} \right) \right|}_{=\lambda} \leq C_T \quad \text{avec } \Delta t, \Delta x \rightarrow 0$$

On cherche alors à quelle condition sur λ l'ensemble

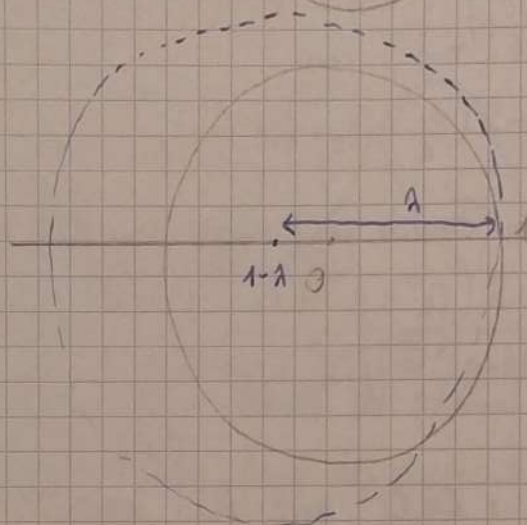
$$1 - \lambda(1 - e^{i\theta}) = \underbrace{1 - \lambda}_{\text{centre}} + \underbrace{\lambda e^{i\theta}}_{\text{rayon}}$$

$\{ 1 - \lambda e^{-ik \frac{2\pi}{L} \Delta x} \}, k \in [-\lfloor \frac{J}{2} \rfloor, \lfloor \frac{J}{2} \rfloor - 1] \}$ et inclus dans $B(0, 1)$ [éventuellement dans $B(0, 1 + \Delta t)$].

Stable



Instable



Le schéma est stable à condition que $1 - \lambda \geq 0$ i.e. $\alpha \frac{\Delta t}{\Delta x} \leq 1$.

Th: Sous condition CFL, le schéma décentré gauche explicite est convergent d'ordre 1 en temps et en espace dans le sens:

$$\sup_{m \in [0, N]} \| (u(t_m, x_j) - u_j^m)_{j \in [0, J-1]} \|_{l^2} \leq C (\Delta t + \Delta x).$$

Preuve: On note $e_j^m = u(t_m, x_j) - u_j^m$.

$$\text{Avec } e_j^{m+1} = e_j^m - (u(t_{m+1}, x_j) - u(t_m, x_j)) + u_j^{m+1} - u_j^m.$$

$$e_j^{n+1} = e_j^n - \Delta t \varepsilon_j^n + \lambda \frac{\Delta t}{\Delta x} (u(t_n, x_j) - u(t_n, x_{j-1}))$$

En utilisant le schéma et l'expression de l'erreur de consistance:

$$\begin{aligned} e_j^{n+1} &= u(t_{n+1}, x_j) - u_j^{n+1} \\ &= u(t_n, x_j) - \lambda (u(t_n, x_j) - u(t_n, x_{j-1})) + \Delta t \varepsilon_j^n - u_j^n \\ &\quad + \lambda (u_j^n - u_{j-1}^n) \end{aligned}$$

$$e_j^{n+1} = e_j^n - \lambda (e_j^n - e_{j-1}^n) + \Delta t \varepsilon_j^n$$

(1) On passe en Fourier pour avoir une belle relation de récurrence:

on m'a pas
supposé u_0
périodique.
 $u_0(0) = u_0(L) \dots$

$$\begin{aligned} C_k^{n+1} &= C_k^n - \lambda (C_k^n - e^{ik \frac{2\pi}{L} \Delta x} C_k^n) + \Delta t F(\varepsilon_j^n) \\ &= (1 - \lambda(1 - e^{ik \frac{2\pi}{L} \Delta x})) C_k^n + \Delta t F(\varepsilon_j^n) \end{aligned}$$

$$e_j^{n+1} = A e_j^n + \Delta t \varepsilon_j^n \text{ avec } A = \begin{bmatrix} 1-\lambda & \lambda \\ 0 & 1-\lambda \end{bmatrix} \quad \begin{matrix} \text{si } u_0 \text{ per.} \\ \lambda^2 \end{matrix}$$

On suppose u^0 (et u) périodique. Alors on peut utiliser Fourier discret:

$$C_k^{n+1} = (1 - \lambda(1 - e^{ik \frac{2\pi}{L} \Delta x})) C_k^n + \Delta t F(\varepsilon_j^n)$$

$$\|C_k^n\|_{\ell^2} \leq \underbrace{\|C_k^0\|_{\ell^2}}_{=0} + \Delta t \sum_{k=0}^{n-1} \|F(\varepsilon_j^k)\|_{\ell^2} \leq C(\Delta t + \Delta x) \quad (\text{car } \|\cdot\|_{\ell^2} \leq \|\cdot\|_{\ell^{\infty}})$$

$$\leq \frac{N \Delta t}{T} C(\Delta t + \Delta x)$$

⑦ Equation de la chaleur 1D Berthelin

On étudie l'équation de la chaleur dans deux situations :

- variable spatiale sur $\mathbb{R}$ tout entier
- variable spatiale dans un segment, avec conditions de bord.

Sur $\mathbb{R}$: $\partial_t u(t, x) = + a \partial_x^2 u(t, x) \quad (t, x) \in]0, +\infty[\times \mathbb{R}$
 (E) $\left\{ \begin{array}{l} u(0, x) = u^0(x) \end{array} \right.$

Analyse de Fourier formelle.

résolution

$$\partial_t \hat{u}(t, \xi) = -a \xi^2 \hat{u}(t, \xi) \quad (\text{EDO à } \xi \in \mathbb{R} \text{ fixé})$$

$$\hookrightarrow \hat{u}(t, \xi) = e^{-a \xi^2 t} \hat{u}_0(\xi)$$

TF inverse

$$\hookrightarrow u(t, x) = \left(\frac{1}{\sqrt{4\pi a t}} e^{-\frac{x^2}{4at}} \right) * u_0(x)$$

Th: On suppose u^0 continue bornée. Alors la fonction :

$$u :]0, +\infty[\times \mathbb{R} \rightarrow \mathbb{R} \\ (t, x) \mapsto \frac{1}{\sqrt{4\pi a t}} \int_{\mathbb{R}} u_0(y) e^{-\frac{(x-y)^2}{4at}} dy$$

est de classe $\mathcal{C}^\infty$, solution de (E), au sens suivant pour la CI :

$$u(t, x) \rightarrow u^0(x^*) \text{ pour tout } x^* \in \mathbb{R} \\ (t, x) \rightarrow (0, x^*)$$

Preuve : • Caractère $\mathcal{C}^\infty$ et solution de l'équation : dérivation sous l'intégrale.

• Condition initiale : chgt de variable + th de continuité des intégrales à paramètre. ▢

On voit sur l'expression de la solution un effet régularisant et dispersant instantané.

Il y a conservation de la positivité, et si u^0 est positive non nulle, $u(t, \cdot) > 0 \forall t > 0$. Toujours pour $u \geq 0$, il y a conservat° de la norme L^1 , et $\|u(t, \cdot)\|_\infty \leq \|u^0\|_\infty$

Sur un segment
$$(E) \begin{cases} \partial_t u(t, x) = a \partial_x^2 u(t, x) & (t, x) \in]0, +\infty[\times]0, L[\\ u(t, 0) = 0, u(t, L) = 0 & \text{(condit° de Dirichlet)} \\ u(0, \cdot) = u^0 \end{cases}$$

Analyse formelle: • Séparer les variables
• Chercher les modes spatiaux compatibles ac cond de bord.

garantit une $\sum \sin(k \cdot)$
Analyse de Fourier: • Symétriser par imparité sur $[-L, L]$.
• u^0 s'écrit comme somme de sa SF $2L$ -périodique (u^0 continue $\mathcal{C}^1$ par morceaux)
• Chercher formellement $u(t, x)$ comme somme de

SF $2L$ -périodique

Th: On suppose $u^0 \in \mathcal{C}^1([0, L])$ avec $u^0(0) = u^0(L) = 0$. On prolonge u^0 en une f^0 $2L$ -périodique impaire et on note
 $u^0(x) = \sum_{k=1}^{+\infty} b_k \sin\left(\frac{k\pi}{L}x\right)$ son développement en série de Fourier.
Alors la fonction $u(t, x) = \sum_{k=1}^{+\infty} b_k e^{-\frac{k^2\pi^2}{L^2}t} \sin\left(\frac{k\pi}{L}x\right)$ est l'unique solution de (E) avec la régularité $u \in \mathcal{C}^\infty([0, +\infty[\times \mathbb{R}) \cap \mathcal{C}^\infty([0, +\infty[\times \mathbb{R})$

⑧ Equation des ondes

On étudie
$$\begin{cases} \partial_t^2 u(t, x) - c^2 \partial_x^2 u(t, x) = 0 & (t, x) \in [0, +\infty[\times \mathbb{R} \\ u(0, \cdot) = u_0, \partial_t u(0, \cdot) = u_1 \end{cases}$$

où $u_0 \in \mathcal{C}^2, u_1 \in \mathcal{C}^1$.

[Comp^t $\neq$ en dim paire & impaire].

On cherche une solution $\mathcal{C}^2$.

Analyse directe: En remarquant $\partial_t^2 - c^2 \partial_x^2 = (\partial_t - c \partial_x)(\partial_t + c \partial_x)$,
on ~~cherche~~ considère $v(t, x) = u(x-ct, x+ct)$.

$$\partial_t v = -c \partial_t u + c \partial_x u$$

$$\partial_t^2 v = -c \partial_t^2 u +$$

on introduit v tq $u(t, x) = v(x+ct, x-ct)$:

$$v(a, b) = u\left(\frac{a-b}{2c}, \frac{a+b}{2}\right)$$

$$\text{Alors } \partial_a v = \frac{1}{2c} \partial_t u + \frac{1}{2} \partial_x u$$

$$\partial_b \partial_a v = -\frac{1}{4c^2} \partial_t^2 u + \frac{1}{4} \partial_x^2 u + \frac{1}{4c} \partial_t \partial_x u - \frac{1}{4c} \partial_t \partial_x u = 0$$

$$\text{Donc } v(a, b) = F(a) + G(b).$$

$$u(t, x) = F(x-ct) + G(x+ct) \text{ qu'on calcule avec les CI.}$$

Analyse de Fourier:

$$\partial_t^2 \hat{u} + (c\xi)^2 \hat{u} = 0$$

$$\text{i.e. } \begin{pmatrix} \hat{u} \\ \partial_t \hat{u} \end{pmatrix}' = \begin{bmatrix} 0 & 1 \\ -(c\xi)^2 & 0 \end{bmatrix} \begin{pmatrix} \hat{u} \\ \partial_t \hat{u} \end{pmatrix}$$

matrice non
diagonalisable.

On résout en distinguant bien $\xi = 0$ et $\xi \neq 0$

On applique les CI et on trouve:

$$\begin{cases} \hat{u}(t, \xi) = \hat{u}_0(\xi) \cos(c\xi t) + \frac{\hat{u}_1(\xi)}{c\xi} \sin(c\xi t) & \xi \neq 0 \\ \hat{u}(t, 0) = \hat{u}_1(0)t + \hat{u}_0(0). \end{cases}$$

cos & sin $\rightarrow$ propag dans les deux sens.

$\times \frac{1}{c\xi} \rightarrow$ primitive.

On peut aussi faire $\partial_t \left(\frac{\partial_x u}{\partial_t u} \right)$ puis Fourier pour avoir des éq. de transport découplées.
sym. en dérivées.

Th. $\forall u_0 \in \mathcal{C}^2, u_1 \in \mathcal{C}^1$, (E) admet une unique solution, $\mathcal{C}^2$,
donnée par $u(t, \cdot) = \frac{u_0(\cdot + ct) + u_0(\cdot - ct)}{2} + \frac{1}{2c} \int_{-ct}^{ct} u_1(\cdot + y) dy$

Propriétés qualitatives

- On a une propagation dans chaque sens à vitesse finie c .
- La régularité de u est donnée par celles de u_0 et u_1 .
- La dynamique est réversible (contrairement à la chaleur).