Réseaux neuronaux informés par la physique pour la conduction thermique avec changement de phase

Bahae-Eddine Madir, Ionut Danaila, Corentin Lothodé, Francky Luddens

Laboratoire de Mathématiques Raphaël Salem, Université de Rouen Normandie

28 mai 2024

Matériaux à changement de phase

Changement de phase dans l'octadécane, production de la convection naturelle dans la partie liquide. (Gong et al. 2015)

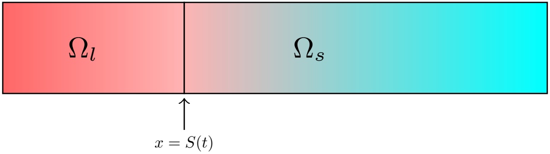
1. Problème de Stefan

L'évolution de la température T dans un matériau à changement de phase est décrite par les équations

$$\partial_t T_l - \alpha_l \partial_x^2 T_l = 0, \qquad \mathcal{T} \times \Omega_l \tag{1a}$$

$$\partial_t T_s - \alpha_s \partial_x^2 T_s = 0, \qquad \mathcal{T} \times \Omega_s \tag{1b}$$

$$k_s \partial_x T_s(t, S(t)) - k_l \partial_x T_l(t, S(t)) = \rho L_f S'(t) \qquad \mathcal{T} \tag{1c}$$



L'enthalpie[1] H du système vérifie l'équation

$$\partial_t H(\theta) - (1/Pe)\,\partial_x^2\theta = 0. \tag{2}$$

En remplaçant dans l'équation, le problème adimensionné peut être écrit comme

$$\partial_t\theta - \frac{1}{Pe}\,\partial_x^2\theta + \frac{1}{Ste}\,\partial_t\left[\varphi_\delta(\theta)\right] = 0 \tag{3a}$$
$$\theta(t_0, x) = g(x), \tag{3b}$$
$$\theta(t, x_l) = \theta_h, \tag{3c}$$
$$\theta(t, x_r) = \theta_c. \tag{3d}$$

En particulier, on cherche à résoudre le problème dans le domaine espace temps $[0,1] \times [0.05, 1]$ lorsque $\delta = 0.05$, $\theta_c = -0.5$, $\theta_h = 1$, $Pe = 100$ et pour les deux nombres de Stefan $Ste = 0.5$ et $Ste = 0.005$.

[1] $H = \theta + (1/Ste)\,\varphi(\theta)$, φ étant la fonction Heaviside
$\varphi_\delta(x) = \frac{1}{2}\left[1 + \tanh(x/\delta)\right]$

Solution de référence

On pose

$$r(\theta) = \frac{1}{Pe}\left(1 + \frac{1}{Ste}\left[\frac{1-\tanh^2(\theta/\delta)}{2\delta}\right]\right)^{-1}, \qquad \delta > 0.$$

On a alors

$$\partial_t \theta = r(\theta)\,\partial_x^2 \theta, \tag{4}$$

le schéma de différences finies avec la discrétisation de Crank–Nicolson associé à l'équation est donné par

$$\frac{\theta_i^{n+1} - \theta_i^n}{\Delta t} = \frac{1}{2}\left[r(\theta_i^{n+1})\frac{\theta_{i+1}^{n+1} - 2\theta_i^{n+1} + \theta_{i-1}^{n+1}}{(\Delta x)^2} + r(\theta_i^n)\frac{\theta_{i+1}^n - 2\theta_i^n + \theta_{i-1}^n}{(\Delta x)^2}\right],$$

l'inconnue θ^{n+1} est la solution d'un système d'équations non linéaires qui peut être résolu par la méthode de Newton-Raphson.

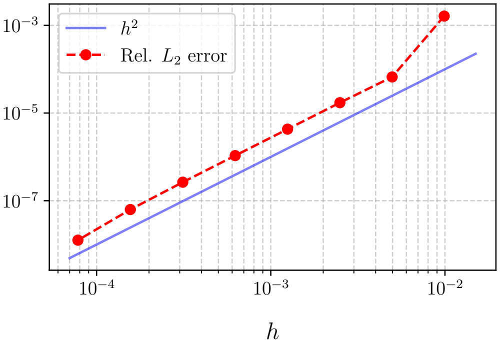


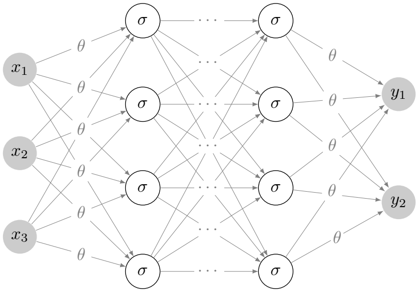
Figure: Convergence quadratique (en maillage) du schéma ($h = \Delta x = \Delta t$).

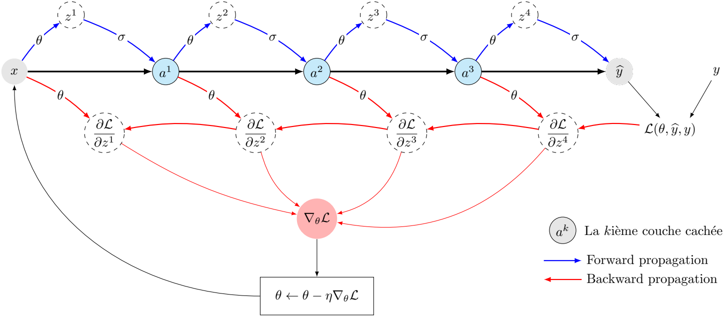
2. Réseaux neuronaux

Un réseau de neurones profond de n couches, est la composition de n fonctions $\ell^k : x \mapsto \ell(x, \theta_k)$, θ_k désigne l'ensemble de paramètres pour la $k$ième couche

$$u_\theta(x) = \ell^n \circ \ell^{n-1} \circ \cdots \circ \ell^1(x). \tag{5}$$

Le plus connu : *multi layer perceptron* $\ell^k(x) = \sigma^k\left(\mathrm{W}^k x + \mathrm{b}^k\right)$



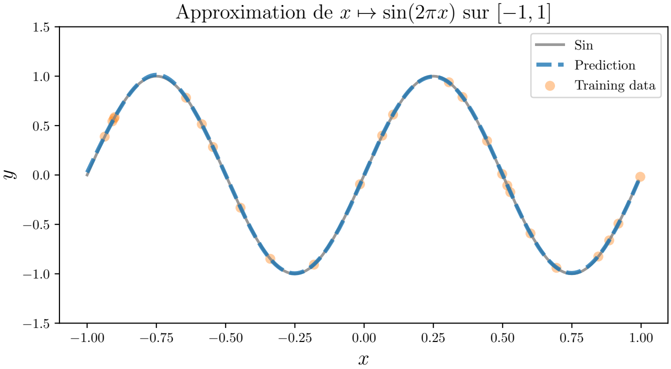


On cherche à résoudre le problème d'optimisation

$$\theta^\star = \arg\min_{\theta} \mathcal{L}(\theta, \widehat{y}, y) \tag{6}$$

$$z^k = \mathrm{W}^k a^{k-1} + \mathrm{b}^k$$
$$a^k = \sigma(z^k) \text{ avec } a^0 = x$$

Un réseau de neurones à une seule couche, peut approcher n'importe quelle fonction continue avec n'importe quelle précision donnée, à condition d'avoir un nombre suffisant de neurones (Hornik. 1991)



3. Réseaux neuronaux informés par la physique

On considère le problème

$$\partial_t u + \mathcal{N}[u] = 0, \qquad \mathcal{T} \times \Omega \tag{7a}$$

$$u(t,x) = g(t,x), \qquad \mathcal{T} \times \partial\Omega \tag{7b}$$

$$u(0,x) = h(x), \qquad \Omega \tag{7c}$$

En remplaçant l'inconnue u par un réseau de neurones $\widehat{u}$, la solution du problème (7) peut être approchée en minimisant les fonctions coûts :

$$\mathcal{L}_r = \|\partial_t \widehat{u} + \mathcal{N}[\widehat{u}]\|_{\mathcal{T}\times\Omega,\, N_r} \tag{8a}$$

$$\mathcal{L}_b = \|\widehat{u} - g\|_{\mathcal{T}\times\partial\Omega,\, N_b} \tag{8b}$$

$$\mathcal{L}_0 = \|\widehat{u} - h\|_{\Omega,\, N_0} \tag{8c}$$

où $\|f\|_{A,\,N} = \frac{1}{N}\sum_{k=1}^{N} |f(x_k)|^2$, les x_k sont choisis de manière uniforme sur A.

Maziar Raissi et al. 2019

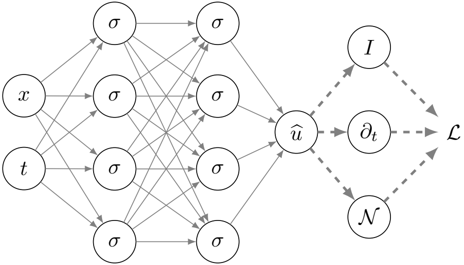


Figure: Physics informed neural network

On considère l'équation de chaleur avec condition de Dirichlet homogène

$$-u''(x) = 4\pi^2 \sin(2\pi x), \qquad x \in [0,1]. \tag{9}$$

Soit $\widehat{u}$ un réseau de neurones d'une seule couche cachée de 10 neurones, et à fonction d'activation $\sigma = \tanh$. On pose

$$\mathcal{L}_r = \frac{1}{N_r} \sum_{k=1}^{N_r} \left(\widehat{u}''(x_r^k) + 4\pi^2 \sin(2\pi x_r^k) \right)^2 \qquad (N_r = 15)$$

$$\mathcal{L}_u = \widehat{u}(0)^2 + \widehat{u}(1)^2$$

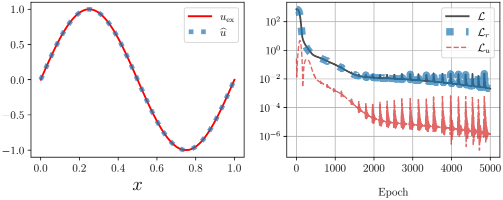


Figure: Minimisation de $\mathcal{L} = \mathcal{L}_r + \mathcal{L}_u$ à l'aide de l'algorithme Adam. Erreur relative L_2 égale à $6.7e-4$.

4. Application sur le problème de Stefan

Résolution par la méthode PINN

hyperparamètres	
Architecture	Deux entrées (t, x) et une sortie $\widehat{\theta}(t, x)$, six couches cachées de 20 neurones à fonction d'activation $\sigma = \tanh$
Initialisation	Xavier
Optimiseur	Adam pour 100000 itérations
Pas d'apprentissage	$t \mapsto \eta\,\gamma^{t/\kappa}$, $\eta = 1e-3$, $\gamma = 0.9$, $\kappa = 2000$
Données d'entraînement	$N_0 = 1024$, $N_b = 256$, $N_r = 10000$ à l'aide de l'échantillonnage Latin Hypercube

$$\mathcal{L}_r = \left\| \widehat{\theta}_t - 0.01\,\widehat{\theta}_{xx} + (1/Ste)\ \left[\varphi_\delta(\widehat{\theta})\right]_t \right\|_{[0.05,1]\times[0,1],\ N_r}$$

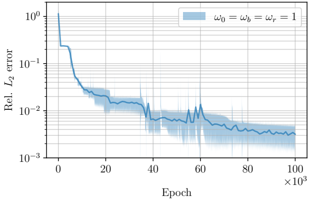
$$\mathcal{L}_0 = \left\| \widehat{\theta}(0.05, \cdot) - g \right\|_{[0,1],\ N_0}$$

$$\mathcal{L}_b = \left\| \widehat{\theta}(\cdot, 0) - \theta_h \right\|_{[0.05,1]} + \left\| \widehat{\theta}(\cdot, 1) - \theta_c \right\|_{[0.05,1],\ N_b}$$

On considère la forme générale de la fonction coût :

$$\mathcal{L} = \omega_0\,\mathcal{L}_0 + \omega_b\,\mathcal{L}_b + \omega_r\,\mathcal{L}_r$$

Cas de $Ste = 0.5$



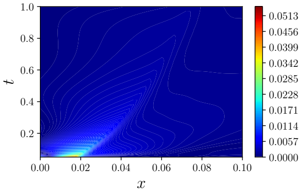


Figure: Erreur relative L_2 (à gauche) de valeur finale $3.1e - 3$. Erreur absolue (à droite).

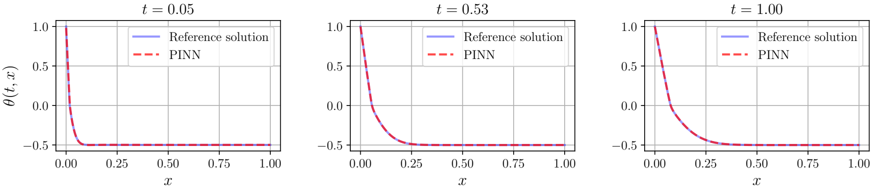
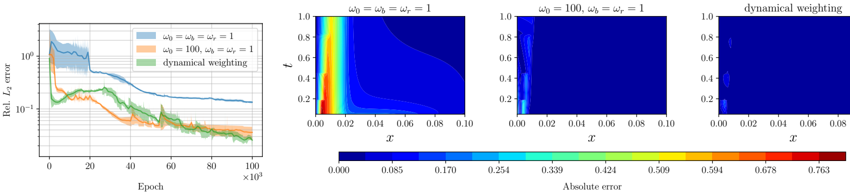


Figure: Solution à la fin de l'entraînement pour $t = 0.05$, $t = 0.53$ et $t = 1$.

Cas de $Ste = 0.005$



Erreurs relatives L_2 en haut: $1.3e-1$ pour le premier cas (blue), $3.5e-2$ pour le deuxième cas (orange) et $2.5e-2$ pour le troisième cas[2] (vert). Erreurs absolues en bas.

[2]dynamical weighting: Sifan Wang et al. 2021.

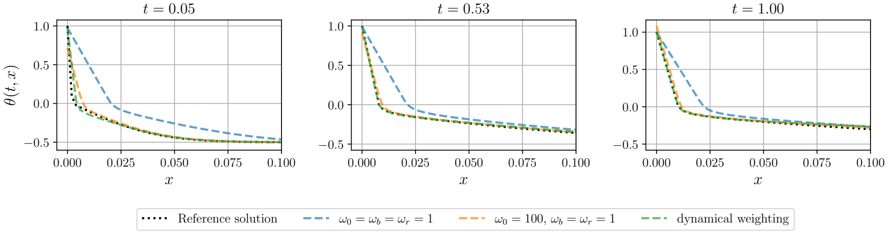


Figure: Solution à la fin de l'entraînement pour $t = 0.05$, $t = 0.53$ et $t = 1$.

Pointwise weighting

Ici les poids w_r, w_0 et w_b dépendent de (t, x)

$$
\begin{aligned}
\mathcal{L}_r &= \left\| m_r(w_r) \left(\widehat{\theta}_t - 0.01\, \widehat{\theta}_{xx} + 200 \left[\varphi_\delta(\widehat{\theta}) \right]_t \right) \right\|_{[0.05,1]\times[0,1],\, N_r} \\
\mathcal{L}_0 &= \left\| m_0(w_0) \left(\widehat{\theta}(0.05, \cdot) - g \right) \right\|_{[0,1],\, N_0} \\
\mathcal{L}_b &= \left\| m_b(w_b) \left(\widehat{\theta}(\cdot, 0) - \theta_h \right) \right\|_{[0.05,1]} + \left\| m_b(w_b) \left(\widehat{\theta}(\cdot, 1) - \theta_c \right) \right\|_{[0.05,1],\, N_b} \\
\mathcal{L} &= \mathcal{L}_r + \mathcal{L}_0 + \mathcal{L}_b
\end{aligned}
$$

m_i, m_b et m_r sont des fonctions positives, différentiables et strictement croissantes. La mise à jour des paramètres se fait par descente de gradient de telle sorte que

$$
\min_{\theta} \max_{\omega_i, \omega_b, \omega_r} \mathcal{L}(\theta, \omega_i, \omega_b, \omega_r) \tag{10}
$$

soit atteint.

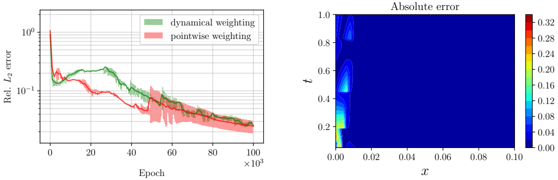


Figure: Erreur relative L_2 (à gauche) de valeur finale $2.4e-2$. Erreur absolue (à droite).

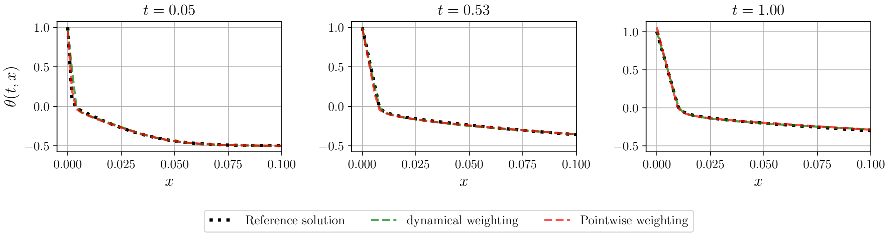


Figure: Solution à la fin de l'entraînement pour $t = 0.05$, $t = 0.53$ et $t = 1$.

5. Conclusion et perspectives

Conclusion et perspectives

- $Ste = 0.5$ on peut directement approcher la solution du problème.
- $Ste = 0.005$ "sharp solution" difficultés dans l'apprentissage, il faut balancer les composantes de la fonction coût, globalement ou localement.
- Couplage du problème de Stefan avec les équations de Navier Stokes (Navier-Stokes-Boussinesq).
- Pour le problème de Stefan, approximation de l'opérateur $(x, t, Ste,\ Pe) \mapsto u(x, t, Ste,\ Pe)$.
- Convergence ?

Merci beaucoup pour votre attention !