

## Cours 2: Séries temporelles, ODE et structures géométriques

---

Emmanuel Franck<sup>\*</sup>,

September 16-20, 2024

**Master CMSI, M2**, Strasbourg

<sup>\*</sup>MACARON project-team, Université de Strasbourg, CNRS, Inria, IRMA, France

The logo for Inria, featuring the word "Inria" in a red, cursive script font.The logo for IRMA, consisting of the letters "IRMA" in a blue, bold, sans-serif font, with a horizontal line underneath. Below the line, the text "Institut de Recherche Mathématique Avancée" is written in a smaller, blue, sans-serif font.

# Outline

---

Séries temporelles et réseaux récurrents

Rappel d'EDO

EDO conservative et structure géométrique

Séries temporelles et réseaux récurrents

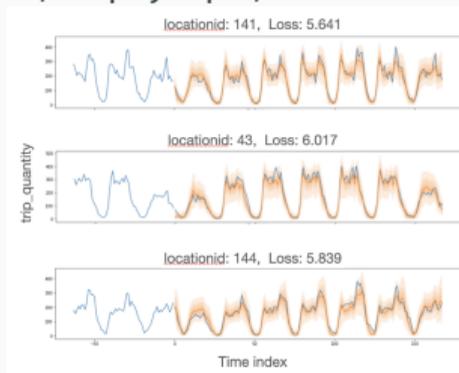
Rappel d'EDO

EDO conservative et structure géométrique

## Séries temporelles et réseaux récurrents

# Séries temporelles et apprentissage

- Un problème classique en apprentissage comme en calcul scientifique est la **prédiction de l'évolution d'une quantité temporelle**.
  - ▶ en calcul scientifique, on part d'un état initial et on prédit l'évolution de cet état,
  - ▶ en apprentissage, on cherche souvent à prolonger une série temporelle déjà entamée. Cela peut s'appliquer en finance, en physique, en traduction etc.



- **Problème de prédiction temporelle:** on connaît la série  $(x_1, \dots, x_n)$  et l'on souhaite construire la loi de probabilité suivante:

$$\mathbb{P}(x_n | x_1, \dots, x_{n-1}), \quad \forall t > 0, \quad \text{ou} \quad \mathbb{E}(x_t | x_1, \dots, x_{n-1})$$

# Réseaux récurrents

- **Approche naive:** utiliser un MLP  $f_{\theta}(x_1, \dots, x_{n-1})$  et prédire  $x_n$ . Problème : un MLP a une taille fixe donc on peut pas le ré-appliquer.
- **Seconde approche:** le **fenêtrage**. On va supposer que le passé lointain compte peu et donc que notre réseau va estimer la probabilité  $\mathbb{P}(x_t | x_{n-1}, \dots, x_{n-m})$ . On peut donc utiliser un MLP.
- **Troisième approche:** le **modèle à variable latente**. On va construire une variable supplémentaire qui va porter au cours du temps de l'information issue des états précédents. On nomme cet état caché  $h_t$  et on cherche donc à construire la loi  $\mathbb{P}(x_n | x_{n-1}, h_{n-1})$ .

## Principe

Un réseau récurrent a pour but de transformer une séquence  $(x_1, \dots, x_n)$  en une autre séquence  $(y_1, \dots, y_n)$  tout en étant capable de traiter de séquence de taille variable. Si vous voulez faire de la prédiction de série temporelle, il suffit de prendre comme valeur de référence pour  $y_n$  l'élément suivant de la séquence  $x_{n+1}$ .

## Définition: réseau récurrent

On se donne des séquences d'entrée  $(x_1, \dots, x_n)$  et de sortie  $(y_1, \dots, y_n)$ . Un réseau récurrent va approximer une fonction de la forme:

$$y_t = f_\theta(x_t, h_{t-1})$$

avec  $x_t$  un élément de notre série temporelle d'entrée,  $h_t$  un vecteur latent et  $y_t$  un élément de notre série de sortie.

## Réseau récurrent de Elman

Soit une entrée  $x_t \in \mathbb{R}^{d_i}$  et une sortie  $y_t \in \mathbb{R}^{d_o}$ . Un **réseau de Elman** est un réseau récurrent à deux couches :

$$\begin{cases} h_t = \sigma_h(W_h x_t + U_h h_{t-1} + b_h) \\ y_t = \sigma_y(W_y h_t + b_y) \end{cases}$$

avec  $d_h$  la dimension des états latents, avec des matrices de poids  $W_h \in \mathcal{M}_{d_h, d_i}$ ,  $U_h \in \mathcal{M}_{d_h, d_h}$  et  $W_y \in \mathcal{M}_{d_o, d_h}$ , des biais  $b_h \in \mathbb{R}^{d_h}$ ,  $b_y \in \mathbb{R}^{d_o}$ .

## Réseaux récurrents profonds

Un réseau récurrent profond à  $K$  couche est réseau du type:

$$\begin{cases} h_t^1 = f_{\theta_h}(x_t, h_{t-1}^1) \\ \vdots \\ h_t^{k-1} = f_{\theta_h}(x_t, h_{t-1}^{k-1}, h_t^{k-2}) \\ h_t^k = f_{\theta_h}(x_t, h_{t-1}^k, h_t^{k-1}) \\ \hat{y}_t = g_{\theta}(h_t^k) \end{cases}$$

avec  $f_{\theta_h}$  une couche du type Elman par exemple et  $g_{\theta}$  un MLP.

## Minimisation et RNN

La fonction de coût pour les réseaux récurrents est donnée par

$$L(\theta_h, \theta) = \frac{1}{T} \sum_{t=1}^T E_t$$

avec  $T$  la taille de la série et  $E_t = (y_t, \hat{y}_t)$ .

# RNN et explosion des gradients

- Gradient de  $E_t$ :

$$\frac{\partial E_t}{\partial \theta_h} = \sum_{i=0}^T \frac{\partial l(y_t, \hat{y}_t)}{\partial \hat{y}_t} \frac{\partial \hat{y}_t}{\partial h_t} \frac{\partial h_t}{\partial h_i} \frac{\partial h_i}{\partial \theta_h},$$

avec  $\theta_h = (W_h, U_h, b_h)$ .

- En appliquant encore une fois la dérivée des fonctions composées:

$$\frac{\partial h_t}{\partial h_i} = \left( \prod_{j=i}^t \frac{\partial h_{j+1}}{\partial h_j} \right) = \prod_{j=i}^t \sigma'_h(W_h x_{j+1} + U_h h_j + b_h) U_h$$

- Ce type de quantité **peut facilement s'annuler pour  $t$  grand** car il suffit que l'un des termes dans le produit s'annule.
- De plus on peut même montrer que cela finira toujours par tendre vers zéro ou l'infini quand la série en temps est longue.
- En effet,  $\lambda_{\max}(U_h) > 1$  cela finira par exploser et si  $\lambda_{\max}(U_h) < 1$  ça finira par disparaître. Cela s'explique par  $|\sigma'_h(W_h x_{j+1} + U_h h_j + b_h)| < 1$  pour "tanh" ou "sigmoïde".

- Plusieurs architectures permettent de limiter ce problème. Par exemple les réseaux **LSTM** et **GRU** (cf cours V. Vigon).

## Cellule LSTM

Soit un état caché  $h_t \in [-1, 1]^d$  et le vecteur d'état  $c_t \in [-1, 1]^d$ . Une cellule LSTM:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o)$$

$$\tilde{c}_t = \sigma_h(W_c x_t + U_c h_{t-1} + b_c)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

$$h_t = o_t \odot \sigma_h(c_t)$$

avec  $W_f, W_i, W_o, W_c \in \mathcal{M}_{d,d}(\mathbb{R})$  les matrices de poids,  $b_f, b_i, b_o, b_c \in \mathcal{M}_{d,d}(\mathbb{R})$  les vecteurs de poids,  $\sigma_g$  la fonction sigmoïde et  $\sigma_h$  la fonction tangente hyperbolique.

Séries temporelles et réseaux récurrents

Rappel d'EDO

EDO conservative et structure géométrique

## Rappel d'EDO

# Apprentissage d'EDO

- L'approche précédente utilise **uniquement un réseau pour approcher la série temporelle**.
- En physique on utilise des **EDOs, résolues numériquement, pour modéliser une évolution temporelle**.

## Approche alternative

Apprendre une ODE et la résoudre numériquement. On peut parler d'approche hybride. Cela peut être plus pertinent les problèmes physiques et permettre de mieux comprendre/étudier la stabilité.

## EDO

On considère une application  $f : I \times U \rightarrow \mathbb{R}^d$  continue localement Lipschitzienne par rapport à  $\mathbf{x}$ . On considère l'EDO:

$$\begin{cases} \mathbf{x}'(t) = f(t, \mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases}$$

## Théorème de Cauchy-Lipschitz

Soit  $I$  un intervalle ouvert de  $\mathbb{R}$ ,  $U$  un ouvert de  $\mathbb{R}^d$  et  $f : I \times U \rightarrow \mathbb{R}^d$  application continue et localement Lipschitzienne en la seconde variable. Alors, si  $t_0 \in I$  et  $\mathbf{x}_0 \in U$  sont donnés, le problème précédent admet une unique solution maximale sur  $I_{max} \subset I$ .

- Le théorème nous donne l'existence et l'unicité localement en temps jusqu'à un certains temps d'existence  $T^*$ .
- Si  $f$  est globalement Lipschitzienne la solution est globale donc  $I_{max} = I$ .
- Le **Lemme des bouts** nous dit que si la solution est **bornée sur**  $[t_0, T^*[$  alors la solution peut être prolongée sur  $\mathbb{R}$ .

# Flot I

- On se replace dans le cadre de l'EDO précédente avec  $I_{max}(t_0, \mathbf{x}_0)$  l'intervalle d'existence de la solution maximale associée aux conditions initiales  $(t_0, \mathbf{x}_0)$ .

## Flot d'une EDO

On appelle **flot local** à l'instant  $t_0$  associé au système  $\mathbf{x}' = f(t, \mathbf{x})$ , l'application  $\varphi^{t_0} : \mathcal{D}^{t_0} \rightarrow \mathbb{R}^d$  telle que  $\varphi^{t_0}(\cdot, \mathbf{x}) : I_{max}^{t_0}(\mathbf{x}) \rightarrow \mathbb{R}^d$  soit la solution maximale de l'équation pour tout  $\mathbf{x} \in U$ .

- Exemple:  $x'(t) = -x(t)$  donc le flot est donné par  $\varphi^{t_0}(t, x) = e^{-(t-t_0)}x$ .
- Le flot est par nature solution de l'EDO.

## Proposition

On suppose  $f \in C^1$  et globalement Lipschitzienne. Alors le flot est défini sur tout  $\mathbb{R} \times \mathbb{R}^d$  et l'application  $t \mapsto \varphi_t(\cdot)$  de  $\mathbb{R} \rightarrow \text{Diff}(\mathbb{R}^d)$  est un homéomorphisme du groupe  $(\mathbb{R}, +)$  dans le groupe  $(\text{Diff}(\mathbb{R}^d), \circ)$  (ensemble des difféomorphismes de  $\mathbb{R}^d$  dans lui-même).

- La proposition précédente est équivalente à dire;
  - ▶  $\forall (s, t) \in \mathbb{R}^2, \boldsymbol{\varphi}_s \circ \boldsymbol{\varphi}_t = \boldsymbol{\varphi}_t \circ \boldsymbol{\varphi}_s = \boldsymbol{\varphi}_{s+t},$
  - ▶  $\boldsymbol{\varphi}_t$  est une bijection,
  - ▶  $(\boldsymbol{\varphi}_t)^{-1} = \boldsymbol{\varphi}_{-t}$  et est continue,
  - ▶  $\boldsymbol{\varphi}_t$  est  $C^1$  d'inverse  $C^1$ .

## Proposition: différentiabilité du flot

Soit  $f$  une fonction de  $D$  dans  $\mathbb{R}^d$  continue et localement Lipschitzienne, telle que  $\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x})$  existe et soit continue sur  $D$ . Alors, le flot est une application continûment différentiable par rapport  $\mathbf{x}$  et sa dérivée  $\Psi_t(\mathbf{x}_0) = \frac{\partial \boldsymbol{\varphi}_t}{\partial \mathbf{x}_0}(\mathbf{x}_0)$  vérifie l'équation:

$$\begin{cases} \dot{\Psi}_t(\mathbf{x}_0) = \frac{\partial f}{\partial \mathbf{x}}(\boldsymbol{\varphi}_t(\mathbf{x}_0))\Psi_t(\mathbf{x}_0) \\ \Psi_0(\mathbf{x}_0) = I_d \end{cases}$$

- Maintenant on va étudier les propriétés géométriques des flots.

### Definition

Supposons  $f$  de classe  $\mathcal{C}^1(D)$ . Elle est de divergence nulle sur  $D$  si

$$\forall \mathbf{x} \in D, \quad \operatorname{div}(f)(\mathbf{x}) = \operatorname{Tr} \left( \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}) \right) = 0$$

### Conservation du volume

Soit un ensemble  $A$ . On définit son volume par

$$\operatorname{Vol}(A) = \int_A d\mathbf{x}$$

Soit une EDO avec  $f$  à divergence nulle alors  $\operatorname{Vol}(\varphi_t(A)) = \operatorname{Vol}(A)$  pour tout ensemble mesurable  $A \subset \mathbb{R}^d$ .

# Schéma numérique I

- Un **schéma numérique est une méthode d'approximation de la solution** d'une équation différentielle. En pratique il va s'agir d'approcher le flot de l'EDO.

## Flot discret ou numérique

Soit  $\Delta t > 0$ . On appelle un **flot discret** que l'on note  $\Phi_{t+\Delta t}^t(\mathbf{x})$  une approximation de  $\varphi^t(t + \Delta t, \mathbf{x})$ .

## Schéma numérique

Soit  $\Delta t > 0$  tel que  $t - t_0 = N\Delta t$  et  $t_n = n\Delta t$  On se donne un flot discret  $\Phi_{t+\Delta t}^t(\mathbf{x})$  une approximation de  $\varphi^t(t + \Delta t, \mathbf{x})$ . Le schéma numérique qui approche  $\varphi^{t_0}(t, \mathbf{x})$  est donné par

$$\Phi_t^{t_0}(\mathbf{x}) = \underbrace{(\Phi_{t_{N-1}+\Delta t}^{t_{N-1}} \circ \dots \circ \Phi_{t_0+\Delta t}^{t_0})}_{n \text{ fois}}(\mathbf{x}).$$

- Les capacités d'approximation des schémas sont reliés aux notions de consistance, convergence et stabilité.

## Consistance

Un flot discret est dit consistant d'ordre  $p$  si

$$\Phi_{t+\Delta t}^t(\mathbf{x}) = \varphi^t(t + \Delta t, \mathbf{x}) + O(\Delta t^{p+1})$$

Le terme  $O(\Delta t^{p+1})$  est appelé l'erreur locale de la méthode.

## Stabilité

Soit  $\Delta t = \frac{T-t_0}{N}$ . Soit une suite  $(\epsilon_n)_{n \in 1, \dots, N}$  avec  $\epsilon_n \in \mathbb{R}^d$ . Un schéma numérique associé à un flot discret  $\Phi_{t+\Delta t}^t(\mathbf{x})$  est dit **stable** par rapport aux erreurs sur l'intervalle  $[t_0, T]$  s'il existe  $\forall N \in \mathcal{N}$  :

$$\|\mathbf{x}_{1,\Delta t}^n - \mathbf{x}_{2,\Delta t}^n\| \leq C \left( \|\mathbf{x}_1^0 - \mathbf{x}_2^0\| + \sum_{m=0}^{n-1} \|\epsilon^m\| \right)$$

avec

$$\mathbf{x}_{1,\Delta t}^{n+1} = \Phi_{t^n+\Delta t}^{t_n}(\mathbf{x}_{1,\Delta t}^n) + \epsilon^n, \quad \mathbf{x}_{2,\Delta t}^{n+1} = \Phi_{t^n+\Delta t}^{t_n}(\mathbf{x}_{2,\Delta t}^n)$$

## Erreur de convergence

Étant donnée une solution  $\mathbf{x}(t) \in C^1([t_0, T]; U)$  on appelle erreur globale:

$$\mathcal{E}(\Delta t) := \max_{0 \leq n \leq T/(\Delta t)} \|\mathbf{x}_n - \mathbf{x}(t^n)\|$$

associé au schéma

$$\mathbf{x}_n = (\Phi_{t_n+\Delta t}^{t_n} \circ \dots \circ \Phi_{t_0+\Delta t}^{t_0})(\mathbf{x}_0).$$

Ce schéma est dit convergent vers  $\mathbf{x}(t)$  sur l'intervalle  $[t_0, T]$  si l'erreur globale tend vers 0 lorsque  $\Delta t$  tend vers 0.

## Théorème de convergence

On suppose l'équation différentielle admet une solution  $\mathbf{x}(t) \in C^1([t_0, T]; U)$ . On considère le schéma associé au flot discret  $\Phi_{t+\Delta t}^t$  avec la donnée initiale  $\mathbf{x}(0) = \mathbf{x}_0$ . S'il est stable sur  $[t_0, T]$  et consistant d'ordre  $p$  avec le flot  $\varphi^{t_0}(t, \cdot)$  de EDO alors il est convergent vers la solution exacte  $\mathbf{x}(t)$ . En outre il existe  $C > 0$  tel que

$$\mathcal{E}(\Delta t) \leq C(\Delta t)^p$$

# Schéma numérique IV

- On rappelle aussi que le flot est solution sur  $[t_n, t_n + \Delta t]$  de

$$\frac{\partial \varphi^{t_n}(t, \mathbf{x}_n)}{\partial t} = f(t, \varphi^{t_n}(t, \mathbf{x}_n)), \quad \varphi^{t_n}(t_n, \mathbf{x}_n) = \mathbf{x}_n$$

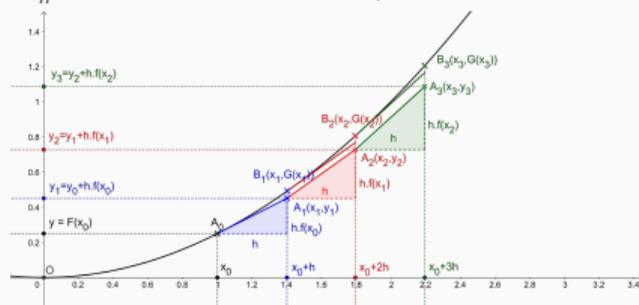
- Si on intègre cette équation on obtient que

$$\int_{t_n}^{t_{n+1}} \frac{\partial \varphi^{t_n}(t, \mathbf{x}_n)}{\partial t} dt = \int_{t_n}^{t_{n+1}} f(t, \varphi^{t_n}(t, \mathbf{x}_n)) dt$$

ce qui donne

$$\varphi^{t_n}(t_{n+1}, \mathbf{x}_n) = \mathbf{x}_n + \Delta t \left( \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} f(t, \varphi^{t_0}(t, \mathbf{x}_n)) dt \right)$$

- ▶ **Euler ex:** méthode des rectangle a gauche,
- ▶ **Euler imp:** méthode des rectangle a droite,
- ▶ **Cranck-Nicolson:** méthode des trapèze,
- ▶ **Runge:** méthode du point milieu + estimateur valeur point milieu par Euler ex.



# Splitting

- On considère l'équation linéaire:

$$\begin{cases} \dot{\mathbf{x}}(t) = A\mathbf{x}(t) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases}$$

que l'on souhaite résoudre sur  $[t_0, t_f]$ . En résolvant l'EDO on obtient que le flot

$$\varphi^{t_0}(t, \mathbf{x}_0) = e^{A(t-t_0)} \mathbf{x}_0 \approx e^{A_2(t-t_0)} e^{A_1(t-t_0)} \mathbf{x}_0$$

avec  $A = A_1 + A_2$ . Est ce que cette approximation est une bonne idée ?

## Erreur de splitting

Soit des matrices  $A, A_1, A_2$  tel que  $A = A_1 + A_2$  L'erreur entre  $e^{At}$  et  $e^{A_1 t} e^{A_2 t}$  est de la forme:

$$e^{At} - e^{A_1 t} e^{A_2 t} = -\frac{t^2}{2} [A_1, A_2] + O(t^3)$$

avec  $[A_1, A_2] = A_1 A_2 - A_2 A_1$ .

- ▶  $e^{At} = I_d + tA + \frac{t^2}{2} A^2 + O(t^3) = I + t(A_1 + A_2) + \frac{t^2}{2} (A_1^2 + A_1 A_2 + A_2 A_1 + A_2^2) + O(t^3)$
- ▶  $e^{A_1 t} e^{A_2 t} = (I + tA_1 + \frac{t^2}{2} A_1^2 + O(t^3))(I + tA_2 + \frac{t^2}{2} A_2^2 + O(t^3)) = I + tA_1 + tA_2 + \frac{t^2}{2} A_1^2 + \frac{t^2}{2} A_2^2 + t^2 A_1 A_2 + O(t^3)$
- ▶ on conclut en faisant la différence

## Splitting II

- Ce résultat nous dit que si on fait cette approximation sur un temps court  $t-t_0$  alors cette approche est pertinente.
- Faire cette approximation est équivalent à résoudre

$$\begin{cases} \dot{\mathbf{x}}_1(t) = A_1\mathbf{x}_1(t) \\ \mathbf{x}_1(t_0) = \mathbf{x}_0 \end{cases}, \quad \text{puis,} \quad \begin{cases} \dot{\mathbf{x}}_2(t) = A_2\mathbf{x}_2(t) \\ \mathbf{x}_2(t_0) = \mathbf{x}_1(t_f) \end{cases}$$

### Schéma de splitting

Soit une EDO de la forme avec  $\mathbf{f}(t, \mathbf{x}) = \mathbf{f}(\mathbf{x}) = \mathbf{f}_1(\mathbf{x}) + \dots + \mathbf{f}_q(\mathbf{x})$  le flot numérique associé au schéma de **splitting de Lie** est de la forme

$$\Phi_{\mathbf{f}, \Delta t}(\mathbf{x}_n) = \Phi_{\mathbf{f}_q, \Delta t} \circ \dots \circ \Phi_{\mathbf{f}_1, \Delta t}$$

avec  $\Phi_{\mathbf{f}_i, \Delta t}$  un flot discret associé à l'EDO portée par  $\mathbf{f}_i$ . Si les flots discrets de chaque équation sont d'ordre 1 ou plus, l'erreur locale de consistance étant en  $O(t^2)$  on a donc un schéma d'ordre 1.

Séries temporelles et réseaux récurrents

Rappel d'EDO

EDO conservative et structure géométrique

## ODE conservative et structure géométrique

# Mécanique Newtonienne

- **La mécanique Newtonienne** est basé sur le point: objet ponctuel idéalisé sur lequel agissent des "forces".
- La position d'un objet données par les coordonnées:  $\mathbf{x}(t) = (x(t), y(t), z(t))$ .
- On peut définir la vitesse ponctuelle  $\mathbf{v}(t) = \frac{d\mathbf{x}(t)}{dt}$  et l'accélération  $\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt}$ .

## Principe fondamentale de la mécanique

Les lois fondamentales de Newton lie l'accélération aux forces extérieures par la relation suivante:

$$m\mathbf{a}(t) = \mathbf{F}(\mathbf{x}(t), \mathbf{v}(t), t) \Leftrightarrow \begin{cases} \frac{d\mathbf{x}(t)}{dt} = \mathbf{v}(t) \\ m \frac{d\mathbf{v}(t)}{dt} = \mathbf{F}(\mathbf{x}(t), \mathbf{v}(t), t) \end{cases}$$

avec la masse  $m$ .

- Exemples de forces:
  - ▶ Le poids:  $\mathbf{F}(\mathbf{x}(t), \mathbf{v}(t)) = -mg\mathbf{k}$
  - ▶ La force de Lorentz pour les particules chargées:  $\mathbf{F}(\mathbf{x}(t), \mathbf{v}(t)) = q(\mathbf{E} + \mathbf{v}(t) \times \mathbf{B})$
  - ▶ La force gravitationnelle de A sur B:  $\mathbf{F}(\mathbf{x}(t), \mathbf{v}(t)) = -g \frac{m_A m_B}{|\mathbf{x}_A - \mathbf{x}_B|^2} \mathbf{u}_{AB}$

# EDO Lagrangienne I

- On va introduire le formalisme **Lagrangien** qui bien que général a d'abord été une reformulation de la mécanique Newtonnienne.
- En général on utilise les positions et les vitesses en coordonnées pour définir la trajectoire d'une particule.
- Dans le formalisme Lagrangien on choisit autre chose. Exemple : pour le pendule on peut prendre l'angle par rapport à l'axe vertical.
- On parle de **coordonnées généralisées**:  $\mathbf{q} \in \mathbb{R}^N$ . On définit aussi **vitesse généralisées** données par  $\dot{\mathbf{q}}$ .

## Principe de moindre action

Formulé par Maupertuis puis Hamilton. Ce principe dit que la dynamique d'une quantité physique (positions, angle, vitesses, etc) peut se déduire à partir d'une unique grandeur appelée **action** en supposant que cette dynamique permettent à l'action d'avoir une valeur optimale entre deux instants donnés.

## Lagrangien

Le Lagrangien est une fonction  $\mathcal{L} : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  dépendante des coordonnées  $\mathbf{q}$  et vitesses généralisées  $\dot{\mathbf{q}}$  qui permet de décrire la dynamique d'un système. Ex:

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{1}{2}m |\dot{\mathbf{q}}|^2 - U(\mathbf{q})$$

avec  $\frac{1}{2}m |\dot{\mathbf{q}}|^2$  l'énergie cinétique et  $U(\mathbf{q})$  l'énergie potentielle.

## Moindre action

Soit un Lagrangien  $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ . Considérons deux points  $\mathbf{x}$  et  $\mathbf{y}$  fixes et deux temps  $t_1$  et  $t_2$ . Soit  $T$  l'espace des courbes paramétrées  $\mathbf{q}(t)$  telles que  $C(t_1) = \mathbf{x}$  et  $C(t_2) = \mathbf{y}$ . **Action:**

$$S(\mathbf{q}(t)) = \int_{t_1}^{t_2} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) dt$$

Le principe de moindre action dit que la dynamique du système est donnée par la courbe  $\mathbf{q}(t)$  minimisant l'action  $S(\mathbf{q}(t))$ .

## Equation d'Euler Lagrange

Soit un Lagrangien  $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$  et  $S$  l'action associée. L'extrema de l'action est solution de l'équation différentielle

$$\frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \dot{\mathbf{q}}} = 0$$

appelée équation de **Euler-Lagrange**.

- On suppose que la trajectoire extrémale existe. Perturbation trajectoire:  $\mathbf{q}_p(t) = \mathbf{q}_s(t) + \alpha \mathbf{f}(t)$ ,  $\dot{\mathbf{q}}_p(t) = \dot{\mathbf{q}}_s(t) + \alpha \dot{\mathbf{f}}(t)$
- Par hypothèse on a  $\mathbf{q}_p(t_1) = \mathbf{q}_s(t_1)$  et  $\mathbf{q}_p(t_2) = \mathbf{q}_s(t_2)$
- La condition d'optimalité est donnée par:

$$\begin{aligned} \frac{dS(\alpha)}{d\alpha} &= \frac{d}{d\alpha} \int_{t_1}^{t_2} \mathcal{L}(\mathbf{q}_p(t), \dot{\mathbf{q}}_p(t), t) dt = \int_{t_1}^{t_2} \frac{d}{d\alpha} \mathcal{L}(\mathbf{q}_s(t) + \alpha \mathbf{f}(t), \dot{\mathbf{q}}_s(t) + \alpha \dot{\mathbf{f}}(t), t) dt = 0 \\ \int_{t_1}^{t_2} \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \mathbf{q}} \frac{d\mathbf{q}_p}{d\alpha} + \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \dot{\mathbf{q}}} \frac{d\dot{\mathbf{q}}_p}{d\alpha} dt &= \int_{t_1}^{t_2} \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \mathbf{q}} \mathbf{f}(t) + \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \dot{\mathbf{q}}} \dot{\mathbf{f}}(t) dt = 0 \end{aligned}$$

- On intègre par partie le second terme et on utilise que  $\mathbf{f}(t_1) = \mathbf{f}(t_2) = 0$  pour avoir

$$\int_{t_1}^{t_2} \left[ \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \dot{\mathbf{q}}} \right] \mathbf{f}(t) dt = 0, \quad \forall \mathbf{f}(t)$$

donc on obtient le résultat.

## Equation du mouvement

$$\left( \frac{\partial^2 \mathcal{L}}{\partial \dot{\mathbf{q}} \partial \dot{\mathbf{q}}} \right) \ddot{\mathbf{q}} + \left( \frac{\partial^2 \mathcal{L}}{\partial \mathbf{q} \partial \dot{\mathbf{q}}} \right) \dot{\mathbf{q}} = \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \mathbf{q}}.$$

- Lagrangien non unique. Exemple: pendule avec  $\theta$  angle vertical. Lagrangien:  
 $\mathcal{L}(\theta, \dot{\theta}) = \frac{1}{2} m l^2 \dot{\theta}^2 + m g l \cos \theta$

## Théorème de Noether

Soit  $\mathbf{f}(\mathbf{q})$  qui paramètre une transformation infinitésimal  $\mathbf{q}_m = \mathbf{q} + \epsilon \mathbf{f}(\mathbf{q})$  qui laisse inchanger le Lagrangien (symétrie) il existe une quantité associée qui est conservée en temps par la dynamique de mouvement. Cette quantité nommée  $\mathcal{Q}$  est donnée par  $\mathcal{Q} = (\mathbf{p} \cdot \mathbf{f}(\mathbf{q}))$  avec  $\mathbf{p} = \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}}$  l'impulsion généralisée.

- les translation spatiale impliquent la conservation de l'impulsion généralisée.
- les rotation impliquent la conservation du moment angulaire  $\mathbf{q} \times \mathbf{p}$ .
- les translations temporelles pour les Lagrangien autonomes impliquent la conservation de l'énergie  $E = (\mathbf{p}, \dot{\mathbf{q}}) - \mathcal{L}$ .
- La structure Lagrangienne intègre les lois de conservation.

# EDO Hamiltonienne I

- On introduit un formalisme alternatif au formalisme Lagrangien, on généralisera et on fera le lien avec la géométrie.
- Un certain nombre de problèmes mécaniques trouvent avantage à être formulés en terme de coordonnées généralisées et d'impulsion généralisée du système.

## Transformée de Laplace

Soit Une fonction  $f(x) : E \rightarrow \mathbb{R}$  une fonction. La transformée de Legendre de  $f$  notée  $f^*(p) : E^* \rightarrow \mathbb{R}$  et appelée fonction conjuguée donnée par

$$f^*(p) = \sup_E (px - f(x))$$

Cette fonction conjuguée est toujours convexe. On en déduit immédiatement l'inégalité de Fenchel  $px \leq f(x) + f^*(p)$ .

Soit un Lagrangien  $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ . Nous définissons **L'Hamiltonien** comme la transformée de Legendre du Lagrangien par rapport à la vitesse généralisée.

# EDO Hamiltonienne II

## Hamiltonien issu d'un Lagrangien

Soit un Lagrangien  $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$  convexe par rapport à  $\dot{\mathbf{q}}$  et coercif. **L'Hamiltonien** s'écrit sous la forme

$$\mathcal{H}(\mathbf{q}, \mathbf{p}, t) = (\mathbf{p}, \dot{\mathbf{q}}) - \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$$

avec  $\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}$  l'impulsion généralisée.

## Système Hamiltonien issu d'un Lagrangien

Soit un Hamiltonien  $\mathcal{H}(\mathbf{q}, \mathbf{p}, t)$  décrivant une dynamique associée au Lagrangien  $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ . Les équations de mouvement sont données par

$$\begin{cases} \frac{d\mathbf{q}}{dt} = \frac{\partial \mathcal{H}(\mathbf{q}, \mathbf{p})}{\partial \mathbf{p}} \\ \frac{d\mathbf{p}}{dt} = -\frac{\partial \mathcal{H}(\mathbf{q}, \mathbf{p})}{\partial \mathbf{q}} \end{cases}$$

Si  $\mathcal{H}(\mathbf{q}, \mathbf{p}, t) = \mathcal{H}(\mathbf{q}, \mathbf{p})$  il est conservé en temps.

- Preuve:

- ▶ On écrit la différentielle de  $\mathcal{H}$ . On obtient

$$d\mathcal{H} = \left( \frac{\partial \mathcal{H}}{\partial \mathbf{q}} d\mathbf{q} + \frac{\partial \mathcal{H}}{\partial \mathbf{p}} d\mathbf{p} \right) + \left( \frac{d\mathcal{H}}{dt} dt \right).$$

- ▶ On applique la même différentielle à la définition du Hamiltonien en fonction du Lagrangien

$$d\mathcal{H} = \mathbf{p} d\dot{\mathbf{q}} + \dot{\mathbf{q}} d\mathbf{p} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} d\mathbf{q} - \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}} + \frac{d\mathcal{H}}{dt} dt$$

- ▶ Puisque  $\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}$  on obtient que

$$d\mathcal{H} = \dot{\mathbf{q}} d\mathbf{p} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} d\mathbf{q} + \frac{d\mathcal{H}}{dt} dt$$

- ▶ On utilise maintenant les équation d'Euler-Lagrange qui disent que  $\frac{\partial \mathcal{L}}{\partial \mathbf{q}} = -\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = -\dot{\mathbf{p}}$ .

- ▶ Si  $\mathcal{H}(\mathbf{q}, \mathbf{p}, t) = \mathcal{H}(\mathbf{q}, \mathbf{p})$

$$d\mathcal{H} = \dot{\mathbf{q}} d\mathbf{p} - \dot{\mathbf{p}} d\mathbf{q}$$

- ▶ Ensuite on remplace la différentielle par les dérivées partielles par rapport à  $\mathbf{p}$  puis  $\mathbf{q}$  on obtient

$$\frac{\partial \mathcal{H}}{\partial \mathbf{p}} = \dot{\mathbf{q}}, \quad \frac{\partial \mathcal{H}}{\partial \mathbf{q}} = -\dot{\mathbf{p}}$$

## Systèmes Hamiltoniens canoniques

On suppose  $\mathbf{x} \in U \subset \mathbb{R}^{2N}$  une EDO Hamiltonienne canonique est de la forme:

$$\dot{\mathbf{x}}(t) = \mathcal{J}^{-1} \nabla_{\mathbf{x}} \mathcal{H}(\mathbf{x})$$

avec

$$\mathcal{J} = \begin{pmatrix} 0 & -I_N \\ I_n & 0 \end{pmatrix}$$

la matrice dit **symplectique**. Le système est dit séparable si il existe  $K : \mathbb{R}^n \rightarrow \mathbb{R}$  et  $U : \mathbb{R}^n \rightarrow \mathbb{R}$  tel que  $\mathcal{H}(\mathbf{q}, \mathbf{p}) = K(\mathbf{p}) + U(\mathbf{q})$ .

## Proposition

Le flot d'un système Hamiltonien canonique préserve le volume

- Preuve: il suffit vérifier que la divergence de  $\mathbf{f} = \mathcal{J}^{-1} \nabla_{\mathbf{x}} \mathcal{H} = 0$ .
- $\nabla \cdot \mathbf{f} = \text{Tr}(\nabla_{\mathbf{x}}(\mathcal{J}^{-1} \nabla_{\mathbf{x}} \mathcal{H})) = \text{Tr}(\mathcal{J}^{-1} \nabla_{\mathbf{x}}^2 \mathcal{H}) = \text{Tr}B - \text{Tr}B = 0$  avec  $B$  les blocs extradiagonaux de  $\nabla_{\mathbf{x}}^2 \mathcal{H}$

## Conservation énergie

Soit  $\phi_t$  le flot associé au système Hamiltonien canonique de Hamiltonien  $\mathcal{H}$ .  $\phi_t$  conserve le Hamiltonien.

- Il suffit de regarder la dérivée temporelle du Hamiltonien.

$$\frac{d}{dt} \mathcal{H}(\phi_t(\mathbf{x})) = \nabla \mathcal{H}(\mathbf{x})^t \dot{\phi}_t(\mathbf{x})$$

puisque le flot satisfait l'équation on a

$$\frac{d}{dt} \mathcal{H}(\phi_t(\mathbf{x})) = \nabla_{\mathbf{x}} \mathcal{H}^t (\mathcal{J}^{-1} \nabla_{\mathbf{x}} \mathcal{H})$$

ce qui donne

$$\frac{d}{dt} \mathcal{H}(\phi_t(\mathbf{x})) = -\nabla_{\mathbf{x}} \mathcal{H}^t (\mathcal{J} \nabla_{\mathbf{x}} \mathcal{H}) = 0$$

par anti-symétrie de  $\mathcal{J}$ .

- Maintenant on va introduire des notions de géométrie symplectique et montrer d'autres propriétés du flot. Ces notions seront très utiles pour l'approximation numérique des EDO Hamiltoniennes.

# EDO Hamiltonienne et géométrie symplectique III

## Forme symplectique dans les espaces vectorielles

Sur un espace vectoriel réel  $V \in \mathbb{R}^N$ , une forme symplectique est une forme bilinéaire  $\omega : V^2 \rightarrow \mathbb{R}$  qui est

- **antisymétrique**: pour tous vecteurs  $\mathbf{v}$  et  $\mathbf{w}$  de  $V$ , on a  $\omega(\mathbf{v}, \mathbf{w}) = -\omega(\mathbf{w}, \mathbf{v})$
- **non dégénérée**: pour tout vecteur  $\mathbf{v} \neq 0$ ,  $\exists \mathbf{w}$  tel que  $\omega(\mathbf{v}, \mathbf{w}) \neq 0$ .

Soit  $\mathbf{x} = (x_1^1, \dots, x_n^1, x_1^2, \dots, x_n^2)$ ,  $\mathbf{y} = (y_1^1, \dots, y_n^1, y_1^2, \dots, y_n^2) \in \mathbb{R}^{2N}$ . La forme bilinéaire

$$\omega(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^N (x_i^1 y_i^2 - x_i^2 y_i^1) = \langle \mathbf{x}, \mathcal{J}\mathbf{y} \rangle_{\mathbb{R}^{2N}}$$

est symplectique et est appelée la forme canonique.

## Symplectomorphisme

Soit une application  $g(\mathbf{x}) : \mathbf{R}^{2N} \rightarrow \mathbf{R}^{2N}$  de Jacobienne  $S(\mathbf{x})$ . Elle est dite **symplectique** si:

$$S^t(\mathbf{x})\mathcal{J}S(\mathbf{x}) = \mathcal{J}$$

## EDO Hamiltonienne et symplectomorphisme

Soit  $\Psi$  un difféomorphisme de  $U \subset \mathbb{R}^{2N}$  dans  $V = \Psi(U)$ . Si  $\Psi$  est une application symplectique le système Hamiltonien s'écrit dans les variables  $\mathbf{y} = \Psi(\mathbf{x})$  sous la forme Hamiltonienne:

$$\dot{\mathbf{y}} = \mathcal{J}^{-1} \nabla_{\mathbf{y}} \mathcal{H}_2$$

avec  $\mathcal{H}_2 = \mathcal{H} \circ \Psi^{-1}$ . Réciproquement si  $\Psi$  transforme une EDO Hamiltonienne en une EDO Hamiltonienne alors c'est une application symplectique.

## Théorème de Poincaré

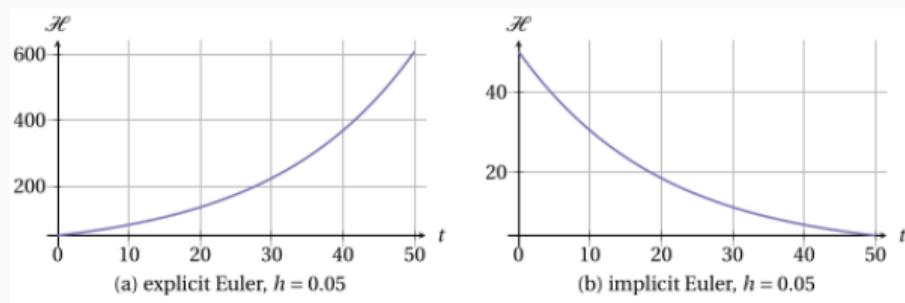
Soit  $\mathcal{H}$  une fonction de classe  $\mathcal{C}^2$  d'un ouvert  $U \subset \mathbb{R}^{2N}$  dans  $\mathbb{R}$  (telle que  $\nabla_{\mathbf{x}} \mathcal{H}$  soit localement Lipschitzienne). Alors pour tout  $(t, \mathbf{x}) \in \Omega_0$ ,  $\varphi^t$  est un **symplectomorphisme**.

- Les **symplectomorphismes** sont essentiels pour les EDO Hamiltoniennes.
  - ▶ Ils laissent invariant la structure Hamiltonienne.
  - ▶ Ils caractérisent les flots d'EDO Hamiltonienne.

# EDO Hamiltonienne et schémas numériques I

- On commence par regarder ce que donne les schémas d'Euler explicites et implicites pour une ODE Hamiltonienne: l'oscillateur harmonique.

$$\mathcal{H}(q, p) = \frac{1}{2}(\omega^2 q^2 + p^2), \quad \begin{cases} \frac{dq}{dt} = p \\ \frac{dp}{dt} = -\omega^2 q \end{cases}$$

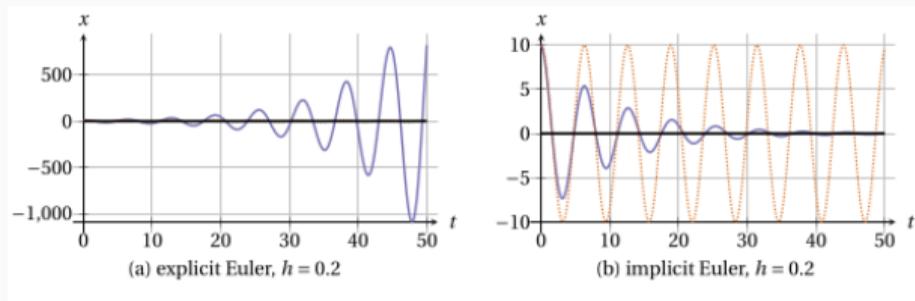


- On démontre que pour Euler explicite  $\mathcal{H}(q^{n+1}, p^{n+1}) = (1 + \omega^2 \Delta t^2) \mathcal{H}(q^n, p^n)$  et  $\mathcal{H}(q_{n+1}, p_{n+1}) = \left(\frac{1}{1 + \omega^2 \Delta t^2}\right) \mathcal{H}(q_n, p_n)$  pour Euler implicite.

# EDO Hamiltonienne et schémas numériques I

- On commence par regarder ce que donne les schémas d'Euler explicites et implicites pour une ODE Hamiltonienne: l'oscillateur harmonique.

$$\mathcal{H}(q, p) = \frac{1}{2}(\omega^2 q^2 + p^2), \quad \begin{cases} \frac{dq}{dt} = p \\ \frac{dp}{dt} = -\omega^2 q \end{cases}$$



- On démontre que pour Euler explicite  $\mathcal{H}(q^{n+1}, p^{n+1}) = (1 + \omega^2 \Delta t^2) \mathcal{H}(q^n, p^n)$  et  $\mathcal{H}(q_{n+1}, p_{n+1}) = \left(\frac{1}{1 + \omega^2 \Delta t^2}\right) \mathcal{H}(q_n, p_n)$  pour Euler implicite.

## Schémas d'Euler

Les flots discrets des schémas d'Euler explicite et implicite **ne sont pas des symplectomorphismes**.

- Preuve:

- ▶ On se restreint au cas séparable. On rappelle le flot

$$\Phi_{\Delta t} : \mathbf{x}_n = \begin{pmatrix} \mathbf{q}_n \\ \mathbf{p}_n \end{pmatrix} \mapsto \begin{pmatrix} \mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n) \\ \mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n) \end{pmatrix}.$$

- ▶ La Jacobienne est donnée par

$$D_{\mathbf{x}_n} \Phi_{\Delta t} = \begin{pmatrix} \nabla_{\mathbf{q}_n} (\mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n)) & \nabla_{\mathbf{p}_n} (\mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n)) \\ \nabla_{\mathbf{q}_n} (\mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n)) & \nabla_{\mathbf{p}_n} (\mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n)) \end{pmatrix}$$

$$D_{\mathbf{x}_n} \Phi_{\Delta t} = \begin{pmatrix} \mathbb{I}_n & \Delta t \nabla_{\mathbf{p}} \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n) \\ -\Delta t \nabla_{\mathbf{q}} \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n) & \mathbb{I}_n \end{pmatrix}.$$

- ▶ Un rapide calcul donne

$$(D_{\mathbf{x}_n} \Phi_{\Delta t})^{\top} \mathcal{J}(D_{\mathbf{x}_n} \Phi_{\Delta t}) = \mathcal{J}(I_n + \Delta t^2 (\nabla_{\mathbf{q}} \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n)) (\nabla_{\mathbf{p}} \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n))).$$

- Le flot discret est symplectique si le système est nilpotent.

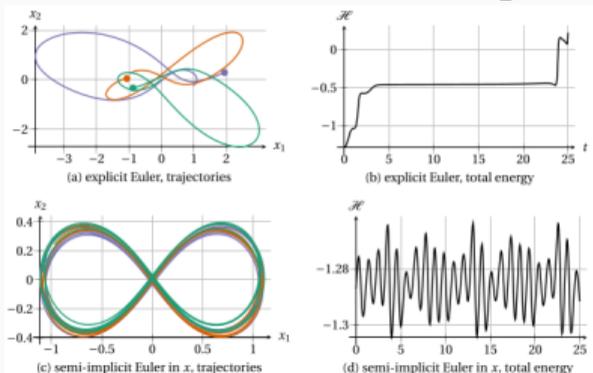
# EDO Hamiltonienne et schémas numériques II

- Idée: construire des flots qui **sont des symplectomorphisme**.

## Schémas d'Euler symplectiques

$$\left\{ \begin{array}{l} \mathbf{q}_{n+1} = \mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{q}_{n+1}, \mathbf{p}_n) \\ \mathbf{p}_{n+1} = \mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_{n+1}, \mathbf{p}_n) \end{array} \right\}, \quad \left\{ \begin{array}{l} \mathbf{p}_{n+1} = \mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n, \mathbf{p}_{n+1}) \\ \mathbf{q}_{n+1} = \mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{q}_n, \mathbf{p}_{n+1}) \end{array} \right\}$$

- Il s'agit de flots discrets semi-implicite d'ordre 1. Explicites dans le cas séparable.
- Flots numériques symplectique préservant l'Hamiltonien modifié:  $\hat{\mathcal{H}}(p, q) = \mathcal{H}(p, q) + \frac{1}{2} \Delta t \omega^2 p q$ . Exemple: 3 corps:



# EDO Hamiltonienne et schémas numériques II

- Preuve:

- ▶ On rappelle le flot

$$\Phi_{\Delta t} : \mathbf{x}_n = \begin{pmatrix} \mathbf{q}_n \\ \mathbf{p}_n \end{pmatrix} \mapsto \begin{pmatrix} \mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n) \\ \mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_{n+1}) \end{pmatrix} = \begin{pmatrix} \mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n) \\ \mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n)) \end{pmatrix}.$$

- ▶ La Jacobienne est donnée par

$$D_{\mathbf{x}_n} \Phi_{\Delta t} = \begin{pmatrix} \nabla_{\mathbf{q}_n}(\mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n)) & \nabla_{\mathbf{p}_n}(\mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n)) \\ \nabla_{\mathbf{q}_n}(\mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n))) & \nabla_{\mathbf{p}_n}(\mathbf{p}_n - \Delta t \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_n + \Delta t \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n))) \end{pmatrix}$$

- ▶ qui se réécrit

$$D_{\mathbf{x}_n} \Phi_{\Delta t} = \begin{pmatrix} \mathbb{I}_n & \Delta t \nabla_{\mathbf{p}} \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n) \\ -\Delta t \nabla_{\mathbf{q}} \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_{n+1}) & \mathbb{I}_n - \Delta t^2 (\nabla_{\mathbf{q}} \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_{n+1})) (\nabla_{\mathbf{p}} \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n)) \end{pmatrix}.$$

- ▶ On a

$$D_{\mathbf{x}_n} \Phi_{\Delta t}^T = \begin{pmatrix} \mathbb{I}_n & -\Delta t \nabla_{\mathbf{q}} \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_{n+1}) \\ \Delta t \nabla_{\mathbf{p}} \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n) & \mathbb{I}_n - \Delta t^2 (\nabla_{\mathbf{q}} \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_{n+1})) (\nabla_{\mathbf{p}} \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n)) \end{pmatrix}$$

et

$$\mathcal{J} D_{\mathbf{x}_n} \Phi_{\Delta t} = \begin{pmatrix} -\Delta t \nabla_{\mathbf{q}} \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_{n+1}) & \mathbb{I}_n - \Delta t^2 (\nabla_{\mathbf{q}} \nabla_{\mathbf{q}} \mathcal{H}(\mathbf{q}_{n+1})) (\nabla_{\mathbf{p}} \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n)) \\ -\mathbb{I}_n & \Delta t \nabla_{\mathbf{p}} \nabla_{\mathbf{p}} \mathcal{H}(\mathbf{p}_n) \end{pmatrix}$$

- ▶ En faisant le produit on obtient  $D_{\mathbf{x}_n} \Phi_{\Delta t}^T \mathcal{J} D_{\mathbf{x}_n} \Phi_{\Delta t} = \mathcal{J}$ .

- Comment on construit les schémas symplectiques dans le cas séparable: on utilise le **splitting**.

$$\left\{ \begin{array}{l} \frac{d\mathbf{q}}{dt} = \nabla_{\mathbf{p}} \mathcal{K}(\mathbf{p}) \\ \frac{d\mathbf{p}}{dt} = 0 \end{array} \right. , \quad \left\{ \begin{array}{l} \frac{d\mathbf{q}}{dt} = 0 \\ \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{q}} \mathcal{U}(\mathbf{q}) \end{array} \right.$$

- donc les flots:

$$\left\{ \begin{array}{l} \frac{d\varphi_{K,q}^{t_n}(t, \mathbf{x}_n)}{dt} = \nabla_{\mathbf{p}} \mathcal{K}(\varphi_{K,p}^{t_n}(t, \mathbf{x}_n)) \\ \frac{d\varphi_{K,p}^{t_n}(t, \mathbf{x}_n)}{dt} = 0 \end{array} \right. , \quad \left\{ \begin{array}{l} \frac{d\varphi_{U,q}^{t_0}(t, \mathbf{x}_n)}{dt} = 0 \\ \frac{d\varphi_{U,p}^{t_0}(t, \mathbf{x}_n)}{dt} = -\nabla_{\mathbf{q}} \mathcal{U}(\varphi_{U,q}^{t_0}(t, \mathbf{x}_n)) \end{array} \right.$$

# EDO Hamiltonienne et schémas numériques IV

- Comment on construit les schémas symplectiques dans le cas séparable: on utilise le **splitting**.

$$\left\{ \begin{array}{l} \frac{d\mathbf{q}}{dt} = \nabla_{\mathbf{p}}\mathcal{K}(\mathbf{p}) \\ \frac{d\mathbf{p}}{dt} = 0 \end{array} \right. , \quad \left\{ \begin{array}{l} \frac{d\mathbf{q}}{dt} = 0 \\ \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{q}}\mathcal{U}(\mathbf{q}) \end{array} \right.$$

- donc les flots:

$$\left\{ \begin{array}{l} \boldsymbol{\varphi}_{K,q}^{t_n}(t_{n+1}, \mathbf{x}_n) = \boldsymbol{\varphi}_{K,q}^{t_{n+1}}(t, \mathbf{x}_n) + \int_{t_n}^{t_{n+1}} \nabla_{\mathbf{p}}\mathcal{K}(\boldsymbol{\varphi}_{K,p}^{t_n}(t, \mathbf{x}_n)) \\ \boldsymbol{\varphi}_{K,p}^{t_n}(t_{n+1}, \mathbf{x}_n) = \boldsymbol{\varphi}_{K,p}^{t_n}(t_n, \mathbf{x}_n) \end{array} \right.$$

et

$$\left\{ \begin{array}{l} \boldsymbol{\varphi}_{U,q}^{t_n}(t_{n+1}, \mathbf{x}_n) = \boldsymbol{\varphi}_{U,q}^{t_n}(t_n, \mathbf{x}_n) \\ \boldsymbol{\varphi}_{U,p}^{t_n}(t_{n+1}, \mathbf{x}_n) = \boldsymbol{\varphi}_{U,p}^{t_n}(t_n, \mathbf{x}_n) - \int_{t_n}^{t_{n+1}} \nabla_{\mathbf{q}}\mathcal{U}(\boldsymbol{\varphi}_{U,q}^{t_n}(t, \mathbf{x}_n)) \end{array} \right.$$

- Pour le premier système le flot associé à  $\mathbf{p}$  reste constant sur l'intervalle  $[t_n, t_{n+1}]$  et idem pour  $\mathbf{q}$  pour le second système. Intégration exacte:

## EDO Hamiltonienne et schémas numériques IV

- Comment on construit les schémas symplectiques dans le cas séparable: on utilise le **splitting**.

$$\left\{ \begin{array}{l} \frac{d\mathbf{q}}{dt} = \nabla_{\mathbf{p}}\mathcal{K}(\mathbf{p}) \\ \frac{d\mathbf{p}}{dt} = 0 \end{array} \right. , \quad \left\{ \begin{array}{l} \frac{d\mathbf{q}}{dt} = 0 \\ \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{q}}\mathcal{U}(\mathbf{q}) \end{array} \right.$$

- donc les flots:

$$\left\{ \begin{array}{l} \boldsymbol{\varphi}_{K,q}^{t_n}(t_{n+1}, \mathbf{x}_n) = \mathbf{q}_{n+1} = \mathbf{q}_n + \Delta t \nabla_{\mathbf{p}}\mathcal{K}(\mathbf{p}_n) \\ \boldsymbol{\varphi}_{K,p}^{t_n}(t_{n+1}, \mathbf{x}_n) = \mathbf{p}_{n+1} = \mathbf{p}_n \end{array} \right.$$

et

$$\left\{ \begin{array}{l} \boldsymbol{\varphi}_{U,q}^{t_n}(t_{n+1}, \mathbf{x}_n) = \mathbf{q}_{n+1} = \mathbf{q}_n \\ \boldsymbol{\varphi}_{U,p}^{t_n}(t_{n+1}, \mathbf{x}_n) = \mathbf{p}_{n+1} = \mathbf{p}_n - \Delta t \nabla_{\mathbf{q}}\mathcal{U}(\mathbf{q}_n) \end{array} \right.$$

- Sous systèmes sont Hamiltonien donc leurs flots exacts sont symplectiques. Pas d'approximation donc les flots discrets aussi.

- Comment on construit les schémas symplectiques dans le cas séparable: on utilise le **splitting**.

$$\left\{ \begin{array}{l} \frac{d\mathbf{q}}{dt} = \nabla_{\mathbf{p}}\mathcal{K}(\mathbf{p}) \\ \frac{d\mathbf{p}}{dt} = 0 \end{array} \right. , \quad \left\{ \begin{array}{l} \frac{d\mathbf{q}}{dt} = 0 \\ \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{q}}\mathcal{U}(\mathbf{q}) \end{array} \right.$$

- donc les flots:

$$\left\{ \begin{array}{l} \mathbf{q}_{n+1} = \mathbf{q}_n + \Delta t \nabla_{\mathbf{p}}\mathcal{K}(\mathbf{p}_n) \\ \mathbf{p}_{n+1} = \mathbf{p}_n - \Delta t \nabla_{\mathbf{q}}\mathcal{U}(\mathbf{q}_{n+1}) \end{array} \right.$$

- **Schéma final**: composition des deux étapes de splitting. La composition de deux symplectomorphisme étant elle même un symplectomorphisme. On a donc un flot numérique **symplectique**.

# EDO Hamiltonienne non canonique

- Les systèmes Hamiltoniens canoniques sont des systèmes dont le flot préserve la forme symplectique canonique
- Cependant on peut aussi définir des modèles qui préservent des formes symplectique non canonique. on parlera de systèmes Hamiltoniens non canoniques.

## Systèmes Hamiltoniens non canoniques

On suppose que  $\mathbf{x} \in U \subset \mathbb{R}^{2N}$ . Une EDO Hamiltonienne noncanonique est de la forme:

$$\dot{\mathbf{x}} = \mathcal{K}(\mathbf{x})^{-1} \nabla_{\mathbf{x}} \mathcal{H}(\mathbf{x})$$

avec  $\mathcal{K}(\mathbf{x})$  une matrice antisymétrique, inversible et satisfaisant l'identité de Jacobi:

$$\sum_{l=1}^n \left( \frac{\partial \mathcal{K}_{ij}(\mathbf{x})}{\partial x_l} \mathcal{K}_{lk}(\mathbf{x}) + \frac{\partial \mathcal{K}_{jk}(\mathbf{x})}{\partial x_l} \mathcal{K}_{li}(\mathbf{x}) + \frac{\partial \mathcal{K}_{ki}(\mathbf{x})}{\partial x_l} \mathcal{K}_{lj}(\mathbf{x}) \right) = 0$$

# EDO Hamiltonienne non canonique II

## K-symplectomorphisme

Soit une forme symplectique donnée par  $\omega(\mathbf{a}, \mathbf{b}) = \langle \mathbf{a}, \mathcal{K}(\mathbf{x})\mathbf{b} \rangle_{\mathbb{R}^{2N}}$  avec  $\mathbf{a}, \mathbf{b}, \mathbf{x} \in \mathbb{R}^{2N}$ . Soit une application  $g(\mathbf{x}) : \mathbb{R}^{2N} \rightarrow \mathbb{R}^{2N}$  de Jacobienne  $S(\mathbf{x})$ . cette application est dite **K-symplectique** si:

$$S^t(\mathbf{x})\mathcal{K}(g(\mathbf{x}))S(\mathbf{x}) = \mathcal{K}(\mathbf{x})$$

## Théorème

Soit  $\mathbf{x} \in U \subset \mathbb{R}^{2N}$  solution de l'EDO non canonique. Soit  $\varphi^{t_0}(t, \mathbf{x})$  le flot associé. Ce flot préserve le volume, préserve  $\mathcal{H}(\mathbf{x})$  et est un **K-symplectomorphisme**.

- Exemple de Lokta Voltera

$$\begin{cases} \frac{dx_1}{dt} = a_{12}x_1x_2 + \lambda_1x_1, \\ \frac{dx_2}{dt} = a_{21}x_1x_2 + \lambda_2x_2 \end{cases}$$

## EDO Hamiltonienne non canonique II

- Exemple de Lokta Voltera sous forme Hamiltonienne -

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & x_1 x_2 \\ -x_1 x_2 & 0 \end{pmatrix} \nabla_{\mathbf{x}} \mathcal{H}(\mathbf{x})$$

avec

$$H(x_1, x_2) = -a_{21}x_1 - \lambda_2 \ln x_1 + a_{12}x_2 + \lambda_1 \ln x_2$$

- Pour obtenir des schémas K-symplectique on utilise souvent le **splitting**.
- Les approches de splitting séparent en général  $\mathcal{K}$  ou  $\mathcal{H}$ .
- Exemple

$$\mathcal{H}(x_1, x_2) = \mathcal{H}_1(x_1) + \mathcal{H}_2(x_2)$$

et

$$\left\{ \begin{array}{l} \frac{dx_1}{dt} = 0 \\ \frac{dx_2}{dt} = a_{21}x_1x_2 + \lambda_2x_2 \end{array} \right\}, \quad \left\{ \begin{array}{l} \frac{dx_1}{dt} = a_{12}x_1x_2 + \lambda_1x_1 \\ \frac{dx_2}{dt} = 0 \end{array} \right.$$

- Les flots de Euler sont exactes pour chaque sous système donc K-symplectique et donc le flot de discret total aussi.

## **Temporary page!**

$\LaTeX$  was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because  $\LaTeX$  now knows how many pages to expect for this document.