Analyse élémentaires de la matrice d'avance et propagateur*

Juin 2005

On traite de divers problèmes soulevés dans l'unsecte concernant l'analyse en éléments propres et singuliers de la matrice d'avance d'état et du propagateur dans le SLTC de modèle dynamique.

1 Avance et sensibilité des éléments propres

On considère une matrice réelle carrée $A(t, \varpi)$ intervenant dans une équation d'avance d'état

$$\partial_t x = A(t)x(t) \tag{1}$$

où $x(t)$ est un écart à la trajectoire de référence $\mathcal{T}(0, T)$. On appelle E la matrice des Vecteurs Propres (V.P.) et Λ la matrice diagonale des valeurs propres (v.p.). De même, F les V.P. à gauche, c'est-à-dire que l'on a par définition :

$$AE = E\Lambda(t) \tag{2}$$

et

$$F^\dagger A = \Lambda F^\dagger(t) \tag{3}$$

On va examiner des propriétés de base de ces concepts pour insister sur un aspect opératoire qu'il me semble important de respecter pour éviter certains court-circuits pouvant mener à des résultats erronés. Il s'agit d'éviter de raisonner ligne par ligne mais de conserver l'aspect matriciel, on verra pourquoi. On montre d'abord le résultat classique de propriétés entre F et E, en multipliant respectivement à gauche et à droite les expressions précédentes par la matrice $F^\dagger$ (et E) ; on obtient l'égalité :

$$F^\dagger AE = F^\dagger E\Lambda = \Lambda F^\dagger E \tag{4}$$

et le produit $F^\dagger E$ commute avec la matrice diagonale Λ, ce qui n'est possible, en supposant tous ses éléments distincts, que si $F^\dagger E$ est elle-même diagonale. Il est alors possible, et on le supposera dans la suite, de normaliser $F^\dagger E = I$. L'un des développement nécessaire de ce travail sera de considérer les cas dégénérés où A possède des v.p. non toutes distinctes, comment utiliser la forme Jordan générale lorsqu'elle n'est pas diagonalisable, etc.

On reprend à présent l'idée de stb d'utiliser cette propriété (4) pour examiner la variation de ces éléments quand le temps change, ou la valeur d'un paramètre,

*.

du SLTC caractérisé par $A(t,\varpi); t \in \{0,T\}$ le long de $\mathcal{T}$. On note ainsi de manière générique par x' une dérivée par rapport à t ou ϖ. La démonstration de stb est la suivante

$$
\begin{aligned}
A\,|e> &= \lambda\,|e> \\
(A\,|e>)' &= A'\,|e> + A\,|e'> = \lambda'\,|e> + \lambda\,|e'>
\end{aligned}
\tag{5}
$$

en considérant un seul élément. Après multiplication par $< f \mid$ associé à $\mid e >$ on obtient successivement

$$
\begin{aligned}
<f\,|\,A'\,|e> + <f\,|\,A\,|e'> &= \lambda' < f \mid e > + \lambda < f \mid e' > \\
<f\,|\,A'\,|e> + \lambda < f \mid e' > &= \lambda' < f \mid e > + \lambda < f \mid e' >
\end{aligned}
$$

d'où le beau résultat

$$
\lambda' = <f\,|\,A'\,|e> \tag{6}
$$

qui permet de connaître la variation d'une v.p., indépendamment de celles des V.P., dès que l'on sait calculer A'[1].

je vous demande de vous arrêter[2]

C'est là que je m'insurge⸮, car ce raisonnement me semble risqué. La dérivation ne prend en effet pas en compte la rotation possible des autres éléments propres de A, on ne prend pas le risque de les considérer. De façon plus générale, il s'agit de la définition même de d_t ou d_ϖ qui est en jeu.

D'où le développement suivant préconisé, qui prend en compte tout l'espace des phases :

$$
\begin{aligned}
(AE)' &= A'E + AE' = E'\Lambda + E\Lambda' \\
F^\dagger(AE)' &= F^\dagger A'E + F^\dagger AE' = F^\dagger E'\Lambda + F^\dagger E\Lambda' \\
&= F^\dagger A'E + \Lambda F^\dagger E' = F^\dagger E'\Lambda + \Lambda'
\end{aligned}
$$

et le résultat

$$
\begin{aligned}
\Lambda' &= F^\dagger A'E + \left[\Lambda, F^\dagger E'\right] \qquad (7) \\
\left[\Lambda, F^\dagger E'\right] &= \Lambda F^\dagger E' - F^\dagger E'\Lambda
\end{aligned}
$$

avec la définition du commutateur $[,]$. Il n'y a aucune raison pour que $F^\dagger E'$ soit diagonale. La propriété liant F a E' provient de la normalisation

$$
(F^\dagger E)' = 0 = F^\dagger E' + F'^\dagger E \tag{8}
$$

Mais par ailleurs les mêmes calculs appliqués à (3) aboutissent au résultat

$$
\Lambda' = F^\dagger A'E + \left[F'^\dagger E, \Lambda\right] \tag{9}
$$

[1] d'où les doubles dérivations de stb et ses élégants **buffers overflow** dans Mortran.

[2] ou *errare humanum est*

qui s'obtient aussi par la substitution de $F^\dagger E'$ par $-F'^\dagger E$ dans (7). Pour être complet, on a encore

$$\left[\Lambda, F^\dagger E'\right] = \Lambda F^\dagger E' - F^\dagger E'\Lambda = \Lambda F^\dagger E' + F'^\dagger E\Lambda$$

A présent, on est en droit de considérer chaque élément de la matrice Λ', qui reste diagonale. Considérant la ligne i et la colonne j du commutateur

$$(\left[\Lambda, F^\dagger E'\right])_{i,j} = (\lambda_i - \lambda_j) < f_i \mid e'_j > \tag{10}$$

on aboutit aux deux résultats suivants, in et hors diagonale :

$$\lambda'_i = < f_i \mid A' \mid e_i > \tag{11}$$
$$< f_i \mid A' \mid e_j > = (\lambda_i - \lambda_j) < f_i \mid e'_j > \tag{12}$$

on peut alors accroire le résultat de stb d'une part, mais aussi remarquer que les $n(n-1)$ équations supplémentaires permettent de trouver les variations de E, ce qui en nécessite n^2. Il reste n contraintes. On peut donc imposer une contrainte sur les modules, soit n. F' se déduit de E'.

2 Le même pour les éléments singuliers

On utilise maintenant la décomposition de A en éléments singuliers, c'est-à-dire que l'on a par définition :

$$\begin{aligned} AV &= U[w]_d(t) \\ U^\dagger A &= [w]_d V^\dagger(t) \end{aligned}$$

où $[w]_d$ est la diagonale définie positive des éléments singuliers, et avec la normalisation

$$UU^\dagger = VV^\dagger = I \tag{13}$$

On répète la dérivation semblable à celle des V.P. :

$$\begin{aligned} (AV)' &= A'V + AV' = U'[w]_d + U[w']_d \\ U^\dagger(AV)' &= U^\dagger A'V + U^\dagger AV' = U^\dagger U'[w]_d + U^\dagger U[w']_d \\ &= U^\dagger A'V + [w]_d V^\dagger V' = U^\dagger U'[w]_d + [w']_d \end{aligned}$$

et le résultat

$$[w']_d = U^\dagger A'V - U^\dagger U'[w]_d + [w]_d V^\dagger V' \tag{14}$$

La normalisation donne de plus

$$(U^\dagger U)' = 0 = U'^\dagger U + U^\dagger U' \tag{15}$$

et ainsi les matrices $U^\dagger U'$ et $V^\dagger V'$ sont antisymétriques. On a aussi le résultat croisé

$$[w']_d = U^\dagger A'V + U'^\dagger U[w]_d - [w]_d V'^\dagger V \tag{16}$$

On peut à présent examiner les termes in et hors diagonale :

$$w'_i = < u_i \mid A' \mid v_i > \tag{17}$$
$$\begin{aligned} < u_i \mid A' \mid v_j > &= w_i < v'_i \mid v_j > - w_j < u'_i \mid u_j > \qquad (18) \\ &= w_i < v'_i \mid v_j > + w_j < u_i \mid u'_j > \end{aligned}$$

du fait de la dite antisymétrie.

3 Application dans Mini_ker

Dans les cas précédents, la matrice A est le tableau dit **aspha** de Mini_ker, qu'il s'agit de dériver par rapport au paramètre choisi par la commande **free-parameter**. En appliquant simplement la règle de dérivation de fonction de fonction à $\partial_t \eta = g(\eta, f(\eta, \varphi), t)$, on trouve, avec $A^{spha} = H + B(I-D)^{-1}C^\dagger$:

$$\begin{aligned} d_\varpi A^{spha} = \quad & d_\varpi H + d_f H d_\varpi f \\ + \quad & [d_f B d_\varpi f + d_\varpi B](I-D)^{-1}C^\dagger \\ + \quad & B[(I-D)^{-1}d_\varpi C^\dagger - (I-D)^{-1}d_\varpi[D](I-D)^{-1}C^\dagger]; \end{aligned} \tag{19}$$

Le calcul des dérivées de matrices – elles mêmes obtenues par dérivation – sont obtenues par double dérivation formelle des fonctions g et f. Ces calculs sont répartis dans trois séquences. Les deux premières sont (**dPi_CtD, dPi_HBbf**), qui donnent les dérivées par rapport au paramètre ϖ des quatre matrices jacobiennes et de la fonction $f(\eta, \varphi$. La troisième séquence (**dffk_HBb**) fournit dfH et d_fB pour la k^{ieme} colonne de la fonction vectorielle f. Un boucle sur $k = 1, m$ permet ainsi de compléter le calcul de la formule (19).

Un utilisation spécifique du type de (11) ou (19) s'effectue dans **ZSTEER**[1].

Algorithme efficace pour la variation des Vecteurs Propres

Il s'agit de calculer E' (et $F'^\dagger$) à partir de la formule (12), qui s'écrit matriciellement

$$\Lambda F^\dagger E' = \Lambda' - F^\dagger A' E + F^\dagger E' \Lambda \tag{20}$$

ou encore

$$E' = PE'\Lambda + Q$$

avec les deux matrices connues : $P = E\Lambda^{-1}F^\dagger; \quad Q = E\Lambda^{-1}[\Lambda' - F^\dagger A' E]$. On résout ce système colonne par colonne (i).

$$|\, e'_i > = (I - \lambda_i P)^{-1} \,|\, q_i >$$

On propose de calculer P^{-1} une fois pour toute et d'appliquer la formule de Woodburry :

$$P \to R^{-1} = P + |\, u >< v \,| \quad \Rightarrow \quad P^{-1} \to P^{-1} - \frac{P^{-1}|\, u >< v \,|P^{-1}}{1 + < v \,|\, P^{-1} \,| u >}$$

Notre application est la plus simple possible, avec $|\, v > = |\, u >$ vecteurs de la base d'Euler de l'espace des états, qui vont sélectionner des lignes et colonnes de P^{-1}, d'où un calcul itératif de $-R = (\lambda_i^{-1} \sum_k |\, u_k >< u_k \,| - P)^{-1}$ avec

$$R^k = R^{k-1} + [\lambda_i(1 + r_{kk}^{k-1})]^{-1} \,|\, r_k^{k-1} >< r_k^{k-1} \,|$$

qui utilise itérativement les $k+1^{eme}$ ligne, colonne, et un terme diagonal de l'itération précédente ayant donnée R^{k-1}. Il reste ensuite à multiplier par la i^{eme} colonne de Q pour avoir $|\, e'_i >$. Le calcul des $F'^\dagger$ s'en déduit par la formule (8).

[1] un exemple est fourni avec **lorhcl**, ou voir avec seb et stb qui sont à l'origine de ce développement.

Pardon, c'est un peu elliptique, je décompose :
on calcule itérativement $-R^n = (\lambda_i^{-1} \sum_k | u_k >< u_k | - P)^{-1}$, avec une première modification ($k = 1$)

$$R_1 = P^{-1} - \lambda_i^{-1} \frac{| p_1^{-1} >< p_1^{-1} |}{1 + p_{11}}$$

en se servant, donc, des premières ligne, colonne, élément diagonal de la P^{-1} originelle (les indices d'itération sont en exposants). La Shermann est ensuite réutilisée pour une modification de R^1 ($k = 2$), avec même formule où cette fois ce sont les deuxièmes[1] ligne, colonne et élément diagonal de R^1 qui sont sollicités :

$$R^2 = R^1 - \lambda_i^{-1} \frac{| r_2^1 >< r_2^1 |}{1 + r_{22}^1}$$

dès qu'on a obtenu R^n, on a $|e_i' >= -R^n \lambda_i \, |p_i>$. Ça va mieux comme ça ?

4 Dynamique des Éléments Propres

Le problème amont[2] est celui de la compréhension des changements structuraux internes des systèmes non-linéaires, à commencer par les É.P. du propagateur : comment, pourquoi, s'effectue une transition paires conjuguées $\leftrightarrow$ réelles du point de vue du SLTC.

Autrement dit, on applique les calculs précédents au cas de la dérivée temporelle et du propagateur $\Phi(t, \tau)$. On a donc (cf 9 avec A matrice d'avance — et ainsi $\partial_t \Phi = A\Phi$) :

$$\begin{aligned} \partial_t \Lambda &= F^\dagger A \Phi E + \left[\Lambda, F^\dagger \partial_t E\right] \\ &= F^\dagger A E \Lambda + \left[\Lambda, F^\dagger \partial_t E\right] \end{aligned} \tag{21}$$

système dans lequel les éléments diagonaux donnent l'avance des v.p. qui se découple heureusement de celle des vecteurs :

$$\partial_t \lambda_i = \lambda_i < f_i | \, A \, |e_i> \tag{22}$$

et où la contrainte donnée par la conservation de la nullité des éléments non-diagonaux va nous permettre d'avancer les V.P. Mais une première remarque s'impose au vu de (22).

Supposant que nous soyons à t du segment (t, τ) de la trajectoire avec des É.P. réels. Clairement dans la formule ci-dessus, rien ne motive un passage en complexes. On est conduit à associer une telle transition à un phénomène de la dynamique qui remette en cause les hypothèses utilisées pour ce calcul. Si on part d'une situation avec une paire complexe conjuguée, on peut affirmer qu'à la transition vers les réels, la dynamique passe par une paire propre de même v.p., et on peut déjà supposer que cela est nécessaire pour que se produise ce changement de corps[3].

Mais finissons le calcul de l'avance dans les mêmes hypothèses que l'obtention de (21). Les termes non-diagonaux donnent :

$$\partial_t \lambda_{ij} \equiv 0 = \lambda_j < f_i | \, A \, |e_j> + (\lambda_i - \lambda_j) < f_i \, | \, \partial_t e_j > \tag{23}$$

[1] remarquez l'utilisation politiquement correcte de l'accord du genre : *Donna prima.*

[2] Al1, Décembre 2006.

[3] *Honni soit qui mal y pense.*

ce qui détermine les $|\partial_t e_j>$. On va essayer de donner une expression par colonne de cette détermination :

$$\begin{aligned} |\sigma_j> &= \lambda_j F^\dagger A\,|e_j> + \Lambda F^\dagger\,|\partial_t e_j> - \lambda_j F^\dagger\,|\partial_t e_j> \\ &= \lambda_j F^\dagger A\,|e_j> + (\Lambda - \lambda_j I)F^\dagger\,|\partial_t e_j> \end{aligned} \tag{24}$$

avec ce vecteur $|\sigma_j>$ formé de zéros sauf en sa composante j qui est formée de $\partial_t \lambda_j = \lambda_j < f_j\,|\,A\,|e_j>$, c'est-à-dire le terme diagonal. Il faut par ailleurs remarquer que les V.P. ne sont déterminés qu'à un facteur près, marqué ici par le fait que la $j$ième ligne de $(\Lambda - \lambda_j I)F^\dagger$ est nulle — dont on notera $|h_j>$ la $j$ième colonne. Le système précédent va ainsi déterminer les composantes de $|\partial_t e_j>$ en fonction de $\partial_t e_{jj}$, que nous prendrons arbitrairement égale à un scalaire non nul α. Pour résoudre ce système, on considère la somme des colonnes de $(\Lambda - \lambda_j I)F^\dagger$ pondérées par les composantes de $|\partial_t e_j>$, qui sont les inconnues sauf α en position j ; il suffit de faire passer cette $\alpha\,|h_j>$ dans les termes connus pour obtenir un système linéaire déterminant l'avance de $|e_j>$ au facteur α près. On peut considérer le système (24) en enlevant sa $j$ième ligne, et les $j$ièmes colonnes des matrices produits[1] :

$$0 = \lambda_j [F^\dagger A]^\flat\,|e_j^\flat> + [(\Lambda - \lambda_j I)F^\dagger]^\flat\,|\partial_t e_j^\flat> + \alpha\,|h_j^\flat>; \tag{25}$$

ce système est régulier avec les hypothèses standards. On a aussi $\partial_t F^\dagger = \partial_t E^{-1}$.

On va vérifier ces formules dans Mini_ker en prenant α donnant aux V.P. la même norme que celle choisie pour le calcul direct sur le propagateur avancé.

De la transition, détours

Que peut-il bien se passer pour qu'on ait une transition réels $\leftrightarrow$ complexes ? Avec une v.p. de multiplicité algébrique 2, qui accompagne nécessairement cette transition, on doit pouvoir pressentir que l'on perd un degré de liberté sur les inconnues λ, sans nécessairement moins de contraintes si les V.P. restent linéairement indépendants, et qu'on en regagne un en passant en complexes ? C'est cette maïeutique (mayonnaise ?) qu'on va essayer de dégager.

Des précédentes hypothèses, on n'a plus $< f_i\,|\,e_j> = 0$ pour i et $j = i+1$ de la paire, et alors $F^\dagger E$ n'est plus diagonale. Par ailleurs, on peut ici faire un parallèle avec la forme réelle canonique associée à une paire. Soit la paire de V.P. $|e> = |e^r> \pm i\,|e^i>$, associée aux v.p. conjuguées $\lambda = \lambda^r \pm i\lambda^i$, avec $M\,|e> = \lambda\,|e>$. On sépare les parties réelles et imaginaires pures :

$$\begin{aligned} M\,|e^r> &= \lambda^r\,|e^r> - \lambda^i\,|e^i> \\ M\,|e^i> &= \lambda^i\,|e^r> + \lambda^r\,|e^i> \end{aligned}$$

que l'on peut écrire symboliquement

$$ME = E \begin{bmatrix} \lambda^r & \lambda^i \\ -\lambda^i & \lambda^r \end{bmatrix} \tag{26}$$

(les deux paires donnent évidemment le même système). On peut ainsi voir un certain lien entre la transition et la non-diagonalité. Cette forme caconique donne ausi la

[1] soit résoudre par la méthode de Householder ?

clé de l'interprètation d'une paire, à savoir que la direction propre est remplacée par un plan propre — dans lequel se déforme en tournant un couple de vecteurs réels.

une technique itérative

Dans ce problème de formulation "physique" de l'avance des éléments propres, comme on pédale dans la choucroute, voici une recette peut-être pas sans rapport[1] donnée par les *Numerical Recipes*. On suppose connue une valeur propre τ proche de λ ; alors, pour un vecteur $|b>$ aléatoire, on approche le V.P. exact $|e>$ en résolvant le système :

$$(M - \tau I)\,|y>=|b> \tag{27}$$

et la v.p. qui s'en déduit

$$\lambda' = \tau + \frac{<b\,|\,b>}{<b\,|\,y>} \tag{28}$$

Pour le voir, on décompose les vecteurs dans la base des V.P. de M :

$$\begin{aligned}(M - \tau I)\sum_j \alpha_j\,|e_j> &= \sum_j \beta_j\,|e_j>\\ \sum_j \alpha_j(\lambda_j - \tau)\,|e_j> &= \sum_j \beta_j\,|e_j>\end{aligned}$$

d'où $\alpha_j = \dfrac{\beta_j}{\lambda_j - \tau}$, et :

$$|y>=\sum_j \frac{\beta_j}{\lambda_j - \tau}\,|e_j> \tag{29}$$

Ainsi, un V.P. "sort" de la somme comme de plus grand module si τ est proche de λ_n, tant que par chance β_n ne se trouve pas trop petit. En itérant la procédure, c'est-à-dire en prenant comme nouveau $|b>\to \dfrac{|y>}{|y|}$, on élève $(\lambda_n - \tau)$ à la puissance deux, etc, et on se rapproche de $|e_n>$.

Pour mettre à jour les v.p., on remarque que

$$\begin{aligned}M\,|e_n> &= \lambda_n\,|e_n> \quad\Rightarrow\\ (M - \tau I)\,|e_n> &= (\lambda_n - \tau)\,|e_n>\end{aligned}$$

en approximant $|e>$ par $|y>$, le terme de gauche est égal à $|b>$ par (27), et on a

$$\begin{aligned}|b> &= (\lambda'_n - \tau)\,|y> \quad\Rightarrow\\ \lambda'_n &= \tau + \frac{<b\,|\,b>}{<b\,|\,y>}\end{aligned} \tag{30}$$

Appliquons cette recette à nos matrices d'avance des È.P. Le propagateur passe en un temps dt de $\Phi(t,\tau)$ à $M\Phi$, avec M approximant e^{dtA}, qui a les mêmes V.P. que A. On part ainsi avec des É.P. $(\lambda, |e>)$ de $M\Phi$ proches de ceux de Φ (soit $\mu, |h>$). La recette donne alors comme première appoximation $|e^1>$ de $|e>$ et λ^1 de λ la solution du système :

$$\begin{aligned}(M\Phi - \mu I)\,|e^1> &= |h>\\ et \quad \lambda^1 &= \mu + \frac{<h\,|\,h>}{<h\,|\,e^1>}\end{aligned} \tag{31}$$

[1] Prenons la comme un pousse-à-idées possible.

Par suivi temporel, la composante β de la projection de $|\,h>$ sur sa suivante $|\,e>$ reste proche de l'unité, ce qui valide la recette. Ce qu'on voit en tout cas, c'est la nécessité d'adjoindre une partie imaginaire non nulle à un $|\,h>$ qui serait réel pur pour suivre une transition.

autre détour

Il reste que la transition dépend de A, une manière de le pointer étant de remarquer que $A\Phi\,|\,e_i>=\lambda_i A\,|\,e_i>$, d'où $<f_j\,|\,A\Phi\,|\,e_i>=\lambda_i<f_j\,|\,A\,|\,e_i>\ \Rightarrow$ c'est A qui dédiagonalise $\Phi(t+dt,\tau)$ dans la base des V.P.

La décomposition univoque $A = T^{-1}JT$ en forme Jordan suggère que la transformation T s'applique aux V.P. ?

On a en effet, pour toute matrice régulière T :

$$\left[T\Phi T^{-1}\right]T\,|h> \;=\; \mu T\,|h> \tag{32}$$

qui rappelle simplement que la similarité matricielle conserve kes v.p. Pour le propagateur mis à jour, on a cette fois-ci :

$$\begin{aligned} A\Phi = T^{-1}JT\Phi\,|e> &= \lambda\,|e> \;\rightarrow \\ J[T\Phi T^{-1}]T\,|e> &= \lambda T\,|e> \end{aligned} \tag{33}$$

Le propagateur ancien, dans la base transformée par A, conserve ses v.p. en héritant des V.P. modifiés. Pour sa mise à jour, la matrice d'évolution A intervient par sa forme Jordan J pouvant dédiagonaliser le précédent propagateur : matriciellement on a

$$\begin{aligned} \left[T\Phi T^{-1}\right]TH &= TH\,[\mu]_d \\ J\;\;\left[T\Phi T^{-1}\right]TE &= TE\,[\lambda]_d \end{aligned} \tag{34}$$

Il semble avec cette remarque que l'on approche d'une microphysique de l'évolution des É.P, non ? On peut aussi la voir comme éloignant Achille du plaisir se rapprochant ; “Achille et la torture”[1] ?

première clôture ?

On considère enfin le cas d'un propagateur $\Phi(t.0)$ diagonalisable à É.P. réels, avec $\Phi\,|\,e>=\lambda\,|\,e>$. L'avance des É.P. est supposée réalisée par $\partial_t[\Phi-\lambda I]=0$. On suppose que E, la matrice des V.P. à droite fournit une base. Il ne semble pas nécessaire de plus d'hypothèses pour proposer la forme suivante :

$$\begin{aligned} (A\Phi-\lambda' I)\,|e> &= -(\Phi-\lambda I)\,|e'> \\ (A\lambda-\lambda' I)\,|e> &= -(\Phi-\lambda I)\,|e'> \end{aligned} \tag{35}$$

oû on a posé $\partial_t\lambda=\lambda'$. On devra revenir sur les hypothèses sous-jacentes à cette première formulation, si on ne veut pas être en contradiction avec le “jvddva” Baladurien. Bien sûr, on ne peut ici résister à la simplification que permet la recherche des vel des propres dans la base E (fixée), c'est-à-dire en posant $|\,e>= E\,|x>$:

$$\begin{aligned} (A\lambda-\lambda' I)\,|e> &= -(\Phi-\lambda I)E\,|x> \\ \lambda_j A\,|e_j> &= -E(\Lambda-\lambda_j I)\,|x_j> +\lambda_i'\,|e_j> \end{aligned} \tag{36}$$

[1] je sens encore comme un reproche de Pat à propos des mouches ...

À l'examen de ce système matriciel, les lignes $j \neq i$ définissent les x_{ij}. Bien sùr, étant donné que la matrice $[\Phi - \lambda_j I]$ est singulière, x_{jj} n'est pas défini, car $x_{jj}(\lambda_j - \lambda_j)$ est nul. Plus extraordinaire, si on remplace ce nul par -1 dans la matrice diagonale $\Lambda - \lambda_j I$, on obtient le terme additionnel $\lambda'_j\,|\,e_j>$!

La résolution du système ainsi formé donne $|\,x_j>$ avec $x_{jj} = \lambda'_j$. On peut alors normaliser $|\,x_j>$.

Passsons à la formulation matricelle du développement précédent ; (35) devient :

$$\begin{aligned} \partial_t(\Phi E - E\Lambda) &= 0 \quad \Rightarrow \\ A\Phi E - E\Lambda' &= -(\Phi E' - E'\Lambda) \\ AE\Lambda &= E'\Lambda - \Phi E' + E\Lambda' \end{aligned} \tag{37}$$

Comme plus haut, on détermine E' dans la base E en posant $E' = EX$:

$$\begin{aligned} AE\Lambda &= EX\Lambda - \Phi EX + E\Lambda' \\ AE\Lambda &= EX\Lambda - E\Lambda X + E\Lambda' \\ E^{-1}AE\Lambda &= X\Lambda - \Lambda X + \Lambda' \\ (E^{-1}AE)_{ij}\lambda_j &= x_{ij}(\lambda_j - \lambda_i) + \lambda'_i\delta_{ij} \end{aligned} \tag{38}$$

Ainsi, ce système permet en effet de résoudre en la matrice $y_{ij} = x_{ij}(\lambda_j - \lambda_i)$ de laquelle on obtient :

$$\begin{aligned} i \neq j : x_{ij} &= \frac{y_{ij}}{\lambda_j - \lambda_i} \\ i = j : \lambda'_i &= y_{ii} \end{aligned} \tag{39}$$

Vérifions cette solution pour $A(t) = C^{\text{ste}}, \quad \Phi(t,0) = exp(At)$. On sait que A et Φ ont mêmes V.P., avec μ et $\exp(\mu t)$ pour v.p. respectives. Alors, $\lambda_j = e^{\mu_j t}$, et $(E^{-1}AE)_{ij}$ est diagonale de terme μ_j. Le système (39) confirme alors que

$$\begin{aligned} i \neq j : x_{ij} &= 0 \\ i = j : \lambda'_i &= \partial_t e^{\mu_i t} = \mu_i\lambda_i = y_{ii} \end{aligned} \tag{40}$$

On a ainsi semble-t-il obtenu un système permettant d'avancer les É.P. dans le cadre des hypothèses posées.

On peut dès lors se reposer la question liminaire du passage $\Re \to \Im$. L'hypothèse que $|\,x>$ soit recherché dans la base E, en supposant qu'elle soit réelle, perd son sens lors d'une transition, mais comment formuler positivement cette impossibilité ?

Une deuxième possibilité consiste à ne pas prendre *a priori* l'hypothèse de diagonalité de Λ' : si en effet on a multiplicité d'une v.p., disons, $\lambda_1 = \lambda_2$; on est contraint d'admettre alors que λ_{12} et λ_{21} ne sont plus nulles. On a vu que ceci devait conduire à une diagonalisation complexe. C'est évident dans le cas où $A(t)$ est antisymétrique et E unitaire, car alors $y_{ij} = -y_{ji}$, ce qui lie les deux lignes de X en une paire conjuguée. Dans le cas général, on peut décomposer la matrice de vel A en ses composantes antisymétrique et symétrique. La composante $\frac{1}{2}(A - A^\dagger)$ prise entre E^{-1} et E devrait complexifier, alors que la partie symétrique en toute intuition devrait amplifier les v.p., mais cette histoire reste à écrire ... peut-être en trouvant la décomposition adequate de A. Une idée ?

vu par le quotient de Rayleigh

Le quotient de Rayleigh associé à la matrice Φ s'écrit $r(x) = \dfrac{<x|\ \Phi\ |x>}{<x\ |\ x>}$. Comme dit Golub à ce sujet, dans le cas où Φ est symétrique, le quotient $r(x)$ est stationnaire lorsque x coïncide avec un V.P. (réel) et alors $r(x) = \lambda$, une v.p. de Φ. On n'a donc plus qu'à différencier en temps avec $\partial_t[.] = \nabla_x[.]\ |x'>$:

$$\begin{aligned}
\lambda &= \frac{<x|\ \Phi\ |x>}{<x\ |\ x>} \\
\lambda' &= \frac{<x|\ A\Phi\ |x'>}{<x\ |\ x>} + 2\frac{<x|\ \Phi\ |x'>}{<x\ |\ x>} - 2r(x)\frac{<x\ |\ x'>}{<x\ |\ x>} \\
\lambda' &= \frac{<x|\ A\Phi\ |x'>}{<x\ |\ x>} + 2\frac{<x|\ [\Phi - \lambda I]\ |x'>}{<x\ |\ x>}
\end{aligned}$$

Le premier terme de la somme donne une contribution à $\partial_t L\lambda$ au plus égal à μ, la v.p. de A du V.P. le plus proche de $|\ e>$. Au contraire, le deuxième terme de la somme est nul lorsque $|\ x]>$ reste aligné avec $|\ x>=|\ e>$, ce qui est correct, puisqu'alors seule l'ampification par λ est modifiée. Ce même terme est nul du fait que Φ est symétrique.

Dans le cas plus général, posons $|x>=|a> + i\,|b>$; on trouve alors que

$$r(x) = \frac{<a|\ A\ |a> + <b|\ A\ |b>}{a^2 + b^2} \tag{41}$$

ce qui conduit à minimiser $\{||(A - \lambda I)\ |a>\ ||_2 + ||(A - \lambda I)\ |b>\ ||_2\}$, et finalement à considérer le quotient de Rayleigh de la partie symétrique de A :
$r(x) = \dfrac{1}{2}\dfrac{<x|\ A + A^\dagger\ |x>}{<x\ |\ x>}$.

5 De la matrice anormale

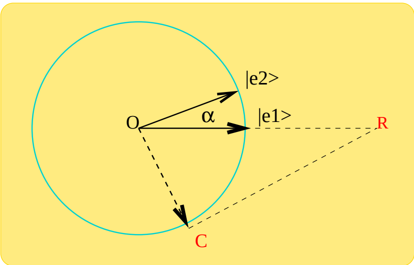
On a vu indépendamment les développements formels auquels donnent accès ces deux décompositions canonique d'une matrice, qui nous intéresse comme analyse de la matrice d'avance ou du propagateur. Ce qui semble moins évident est d'une part, le passage de l'une à l'autre, et d'autre part leurs avantages respectifs. On va essayer de le comprendre en suivant une analyse de Farell et Yoannou (F&Y)[1] qui depuis une quinzaine d'années, d'article en article, mettent en évidence les conséquences dynamiques de la non-normalité de la matrice d'avance.

Leur première remarque concerne l'analyse de stabilité par décomposition modale en ce qu'elle est perverse. La propriété fondamentale d'une matrice normale est que ses vecteurs propres forment une base orthogonale. F&Y déclinent diverses conséquences de non-orthogonalité de cette base, en commençant par l'analyse de stabilité.

On étudie la croissance d'une perturbation initiale : $\sigma^2(t) = \dfrac{< x(t) \mid x(t) >}{< x_0 \mid x_0 >}$, en faisant la remarque que l'on peut toujours se ramener à une norme euclidienne si on lui a préféré a priori une norme $< x \mid M \mid x >$ avec M symétrique (ou hermitienne) définie positive. Alors $\sigma^2(t) = \dfrac{< e^{A^\dagger t} e^{At} x_0 \mid x_0 >}{< x_0 \mid x_0 >}$ pour un système autonome, si bien que la croissance la plus forte sera liée au mode $\lambda_{max}(e^{A^\dagger t} e^{At})$, réel par symétrie.

Si à présent on considère la SVD du propagateur : $e^{At} = U\sigma V^\dagger(t)$, alors $e^{A^\dagger t} e^{At} = V\sigma^2 V^\dagger(t)$ et la croissance est encadrée : $\sigma^2_{min} \leq e^{(\lambda_i + \bar{\lambda}_i)t} \leq \sigma^2_{max}$, où λ_i parcourt les modes de A. On a ainsi deux visions de croissance, la vision modale qui explicite la croissance exponentielle dùe au premier mode, et d'autre part à chaque instant, un vecteur $\mid v_1(t) >$ perturbation optimale du système à cet instant, causant un effet optimal le long de $\mid u_1(t) >$ par multiplication par σ_1.

Si maintenant on considère le mode dominant $e^{\lambda_1 t} \mid e_1 >< e_1^{-1} \mid$, l'excitation max du V.P dominant (supposons le réel) se fait le long du vecteur dont la projection sur $\mid e_1 >$ est la plus grande. Avec l'opérateur de projection $\mid e_1 >< e_1^{-1} \mid$, qui se fait parallèlement au vecteur perpendiculaire à $\mid e^{-1} >$. Quand la base des V.P. est orthogonale (matrice normale) cela coïncide avec le vecteur lui-même, ce qui n'est plus le cas dès que la base n'est plus orthogonale (matrice non-normale). On peut le constater par l'apparition de $E^\dagger E$ dans $\sigma^2(t) = E^{-\dagger} e^{\bar{\lambda} t} E^\dagger E e^{\lambda t} E^{-1}$. Celà se conçoit



Excitation optimale d'un V.P.

Le vecteur excitant le mieux $\mid e_1 >$ est son binormal, orthogonal au deuxième V.P. de la base en 2D : c'est $\overrightarrow{OC}$, le long de $\mid e_1^{-1} >$[a]. Le résultat est $\lambda_1 \overrightarrow{OR}$, et est d'autant plus grand que l'angle est faible (augmentant l'anormalité).

[a] l'excitation optimale est définie comme un vecteur de norme unité.

facilement en deux dimensions comme dans la figure cî-dessus. On peut se convaincre alors que $\mid v_1 >=\mid e_1^{-1} >, \mid u_1 >=\mid e_1 >$ (normés à 1) et que $w_1 = e^{\lambda_1 t} \csc(\alpha)^2$ (le sens

[1] *in* JAS 1996, Vol. **53**, No. 14 pp. 2025–2040.
[2] Ceci est à retenir et généraliser pour formaliser le lien entre les deux décomp.

de $\overrightarrow{OC}$ étant choisi pour donner un $w_1 > 0$).

limite asymptotique. Lorsque $t \to \infty$, l'évolution s'aligne bien sur le mode dominant ($e^{At} \to e^{\lambda_1 t} | e_1 >< e_1^{-1} |$), mais il reste que l'excitation (initiale) optimale est donnée par la binormale à $| e_1 >$, et non ce vecteur lui même, ce qui est moins reconnu si l'on croit F&Y.

limite aux temps courts. Comme on l'a déjà vu dans les Cahiers (SLTC), un développement de Taylor de $\sigma^2(t)$ montre que les E.P. de $A + A^\dagger$ déterminent la croissance max, et diverses études d'écoulement ont montré que ces bouffées de croissances peuvent se produire alors même que tous les modes normaux de A soient déjà amortis.

Implication sur la dynamique stochastique ...

Soit le système

$$d_t x = Ax + Bw(t)$$

où w est un bruit blanc ($\overline{w_i(t)w_j(s)} = \delta_{ij}\delta(t-s)$, avec une moyenne d'ensemble. B réparti les sources, et dès qu'elle est prise comme unitaire, la stat. n'en dépend plus. On a alors

$$\overline{\|x(t)\|^2} = TrB^\dagger \int_0^t e^{A^\dagger(t-\tau)} e^{A(t-\tau)} d\tau \, B = TrB^\dagger Q B$$

Ainsi, l'accumulation de croissance par perturbation plate est révélée par l'opérateur symétrique $Q^t = \int_0^t e^{A^\dagger \tau} e^{A\tau} d\tau$, que l'on peut obtenir par intégration du système

$$d_t Q^t = A^\dagger Q^t + Q^t A + I$$

Lorsque A est asymptotiquement stable, le système tend vers un état d'équilibre statistique, et obéit à la relation de Lyapunov :

$$A^\dagger Q^\infty + Q^\infty A = -I$$

On obtient alors les forçages stochastiques optimals (à mettre en colonne de B) par analyse de Q^∞, hiérarchisés par leur contribution à la variance de l'état. On voit l'analogie avec la précédente analyse de croissance optimale, la matrice Q^t est à présent l'intégrale de $e^{A^\dagger t} e^{At}$.

On a déterminé ce qui conserve le mieux la variance, mais pas comment elle se distribue au sein de l'espace des états. Ces structures orthogonales qui portent la variance sont les EOF (on peut voir une analogie avec les $| v >$ qui excitent le mieux le système d'une part, et les $| u >$ qui donnent la structure de cette variance. On obtient les EOF à partir de la matrice de corrélation :

$$P^t = \overline{| x >< x |} = \int_0^t e^{A(t-\tau)} BB^\dagger e^{A^\dagger(t-\tau)} d\tau$$

qui satisfait l'équation d'évolution

$$d_t P^t = AP^t + P^t A^\dagger + BB^\dagger$$

avec les mêmes hypothèses de stabilité, on a aussi :

$$AP^\infty + P^\infty A^\dagger = -BB^\dagger$$

Chaque v.p. de P^t explique la part de variance contenue dans la structure donnée par un V.P. à chaque instant t. Ceci est donc *a priori* différent du choix de la structure initiale excitant le plus le système. Il s'avère que pour les matrices normales (et avec $BB^\dagger = 1$), ces deux structures coïncident (toutes les matrices commutent) : alors, les modes normaux, les EOF et les optimales stochastiques sont identiques. Ce n'est plus le cas en contexte d'anormalité : on a alors d'une part des colonnes de B (B unitaire, comme V) qui contribuent hiérarchiquement à la variance maintenue, et de l'autre, la répartition de cette variance entre structures d'état (comme U).

... et aspects spectraux

En remplaçant le forçage stochastique Bw par une source $f(t)$, la réponse en transformée de Fourier est $\hat{x}(\omega) = R(\omega)\hat{f}(\omega)$, avec $R(\omega) = (i\omega - A)^{-1} = E(iw-\lambda)^{-1}E^{-1}$, avec les E.P. de A, et l'inégalité de Schwartz montre que

$$\|R(\omega)\| \leq \frac{\kappa(E)}{dist_{min}(iw, \Lambda(A))}$$

où $\kappa(E) = \|E\|\|E^{-1}\|$ est le conditionnement de E, qui devient égal à l'unité dans la cas normal[1]. Cela indique donc la possibilité d'un fort accroissement de variance par rapport au cas normal où seule intervient la distance minimum d'une fréquence donnée $(i\omega)$ aux v.p. de A de son spectre $\Lambda(A)$:

$$\|R(\omega)\| = \frac{1}{dist_{min}(iw, \Lambda(A))}$$

Ainsi, dans le cas normal, le forçage optimal se fait à la fréquence ω de la v.p. la moins amortissante du spectre de A et la variance décroît avec le carré de son exponentielle. Dans le cas de non-normalité, la fréquence de réponse max n'est plus strictement celle de la v.p. dominante, et l'amplitude peut être beaucoup plus forte.

Lors que toutes les fréquences sont également excitées — re-bruit blanc de variance unitaire : $< \hat{f}_i(\omega_1 \mid \hat{f}_j(\omega_2) > = \dfrac{\delta_{ij}\delta(\omega_1-\omega_2)}{2\pi}$, la moyenne d'ensemble de la variance est

$$< x^2 >= \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)d\omega$$

avec $F(\omega) = TrR^\dagger(\omega)R(\omega)$. Comme les v.p. de R sont $(iw-\lambda_i)^{-1}$, on a

$$F(\omega) \leq \sum_{i=1}^{n} \frac{1}{|iw-\lambda_i|^2}$$

avec égalité dans le seul cas normal (Perceval-Planchoreille). Le cas non-normal n'interdit pas qu'une interaction entre les modes normaux permette de tirer du fond bruité l'excitation d'une bande de fréquence particulière.

[1] E peut alors être choisi unitaire.

La morale de tout ceci semble être que l'analyse modale n'est pas la réponse à tout — du moins dans les cas de non-normalité de la matrice d'avance, et ceci parce que l'on ne dispose alors pas de base propre orthonormée, alors que la SVD en fournit, mais au prix il est vrai de l'abandon de la fonction temporelle explicite. C'est le message de F&Y, et on verra en décomposition de domaine que la réduction par hiérarchie des modes normaux est une réduction peu efficace.

6 Vers la conjecture de Goodman et Marshall

La question est : *qu'est-ce qu'un système inerte ?* Et dès lors que l'on dispose des précédentes décompositions, comment varient les éléments propres et singuliers ? Au-delà, comment définir un système réduit lent répondant à un forçage stochastique non linéairement filtré[1] ?

Intuitivement, si le système "évolue lentement[2]", on aimerait formuler cette hypothèse, et en déduire que les v.p. évoluent moins vite que l'état du système lui-même, de même que les vecteurs associés. C'est une première étape. La seconde devra alors prendre en compte l'effet de la partie rapide du système sur la lente, ce qui devrait nous conduire à la dite conjecture ?

Et comme après un premier jet je m'aperçois que la confusion – dans ma tête en tout cas – règne entre les notions de lenteur, d'inertie, et d'autonomie du système, je reprends calmement...

6.1 intro programmatique $[[\sigma]]$

La première hypothèse concernant le temps est en réalité contenue dans les hypothèses plus ou moins implicites que l'on fait en appliquant le TEF pour calculer une trajectoire : on fait porter la validité de l'approximation sur le choix de l'incrément temporel δt, en développant l'avance à l'ordre trois[3]

$$\delta\eta = g_0\delta t + Ag_0\frac{\delta t^2}{2} + (A^2 g_0 + \partial_t A\, g_0)\frac{\delta t^3}{3!} + \mathcal{O}(\delta t^4)$$

et en imposant

$$\delta t < \tau : \quad \|(A^2 + \partial_t A)g_0\|\frac{\tau^3}{3!} \ll \|Ag_0\|\frac{\tau^2}{2}$$

soit donc

$$\delta t \ll \tau : \quad \tau \simeq 3\|A + \partial_t A\, A^{-1}\|^{-1} < \frac{3\|Ag_0\|}{\|(A^2 + \partial_t A)g_0\|} \tag{42}$$

on trouve classiquement qu'il suffit de contraindre l'incrément temporel par les constantes de temps de A dans le cas autonome[4], à un facteur près[5]. Le précédent développement est purement numérique mais rejoint le TEF par l'ordre d'approximation (deux ici comme pour le TEF). On peut aussi remarquer que le terme non-autonome $\partial_t A$ n'apparaît qu'au troisième ordre en opposition avec A^2, ce qui fait que **le TEF épuise la précision du modèle** à l'ordre deux en temps.

Il m'apparaît de manière plus claire que la confusion susdite est liée à la nature du problème que l'on se pose, qui n'est pas classiquement celui de la réponse inerte du système à une source donnée, mais de sa réponse à une perturbation. On est donc de nouveau dans le contexte SLTC et le problème se complique : la matrice $A(t)$ peut varier le long de $\mathcal{T}^{\mathrm{ref}}$. Alors, quel sens donner à l'inertie d'un système non-autonome ? En fait, on soulève à la fois un problème d'inertie du SLTC et de

[1] c'est une interprétation forte de la conjecture.

[2] en référence à une décomposition lent-rapide d'un système primitif.

[3] $\partial_t\eta = g(\eta)$: pas de fonction explicite du temps, $A = \partial_\eta g$, et ainsi $\partial_t g = A\partial_t\eta = Ag$.

[4] non, pas autonome, qui signifie que le système n'a pas de terme explcite en temps $d_t x = f(x,t)$, comme nous ont fait remarqué un reviewer ...

[5] on peut encore ici noter l'importance du chiffre 3 dans le TEF, qui dépasse celui du genre.

la stabilité du système : il pourrait lentement être instable, c'est-à-dire réagir de manière inerte à une perturbation (à la limite un bruit) en restant stable ou non !

Pour essai, on va d'abord attaquer la quasi-autonomicité et tenter une séparation de $A(t) = \bar{A} + N(t)$ entre partie autonome et non-A (pour les fans de Van Vogh).

On a posé $\bar{A} = \dfrac{1}{T}\displaystyle\int_0^T A(t)dt$, et par voie de conséquence $\displaystyle\int_0^T N(t)dt = 0$[1]. La question est d'examiner une séparation possible – à la Floquet – du propagateur :

$$\partial_t \Phi(t,s) = \bar{A}\Phi + N(t)\Phi \rightsquigarrow \Phi = \exp \bar{A}(t-s)\, \Psi(t,s) \quad ? \tag{43}$$

on cherche ainsi à déterminer Ψ via $B(t)$ telle que $\partial_t \Psi = B\Psi$:

$$\begin{aligned} \partial_t \Phi &\overset{A}{=} \bar{A}\,\Phi(t,s) + N(t)\exp \bar{A}(t-s)\Psi(t,s) \\ &\overset{(43)'}{=} \bar{A}\,\Phi(t,s) + \exp \bar{A}(t-s)\, B(t)\Psi(t,s) \end{aligned}$$

d'où en résumé le résultat sur le propagateur de la décomposition :

$$\begin{gathered} \begin{cases} \Phi(t,s) = e^{\bar{A}(t-s)}\Psi(t,s) \\ \partial_t \Psi(t,s) = B(t; t-s)\,\Psi(t,s) \end{cases} \\ B(t;t-s) = e^{-\bar{A}(t-s)}\, N(t)\, e^{\bar{A}(t-s)} \end{gathered} \tag{44}$$

ainsi, Ψ fonctionne différemment de Φ, car B dépend aussi de $(t-s)$ — comme pour Floquet d'ailleurs. On peut remarquer que B a les mêmes v.p. que N (et que $S^{-1}NS$).

La surprise[2], c'est que si $\bar{A}$ et $N(t)$ ne commutent pas, l'évolution de Ψ est amplifiée par le conditionnement de $\bar{A}$. Plus explicitement, dès qu'une perturbation est alignée avec le V.S. optimal ($\mid v_1 >$) de $N(t)$, elle est multipliée par sa v.s. optimale w_1, et de plus potentiellement par $\frac{|\lambda_{max}|}{|\lambda_{min}|}$ conditionnement de $\bar{A}$ dès que $[\bar{A}, N] \neq 0$ (il suffit qu'elle rentre par le max à droite de N et sorte par le min à gauche).

Rien n'empêchera donc un système quasi-autonome, lui-même stable, de ne pas sortir de sa quasi-autonomie – qui porte sur $\|N(t)\|$ – par un double effet de non-commutation et de largeur de spectre.

Remarque pour les Y de l'unsecte que les X avaient immédiatement notée : notre petite séparation $\bar{A} + N(t)$ nous ramène pour la définition de l'inertie au problème précédent. On peut légitimement faire porter sur $\bar{A}$ la caractérisation de l'inertie du système. La comparaison des normes de $N(t)$ et de $\bar{A}$ permettra de définir sa quasi-autonomicité. Quant à sa stabilité, avec ce qu'on en a déjà dit, on pourra ou non la prendre en compte pour rendre plus contraignante la définition de l'inertie – on pourrait y introduire le rayon spectral de $\bar{A}$, ou encore imposer que l'effet des v.p. dépassant celles de $\bar{A}$ soit négligeable. **on a ainsi à définir des concepts dans la couche physique, et je sollicite à nouveau vos contributions**.

[1] al1, 22/04/09, Cécile Penland prend des moyennes d'ensemble pour obtenir $\exp(\bar{A}\tau) = \overline{\Phi(t+\tau,t)}$, avec encore assimilation à une var-covar stationnaire : $\overline{\mid x(t+\tau) >< x(t) \mid}\;\overline{\mid x(t) >< x(t) \mid}^{-1}$.

[2] pour les non-lecteurs de Farell&Ioannou qui exhibent le rôle de la non-normalité des matrices en tant que mère d'instabilités vicieuses – si tant est que la matrice-mère puisse être normale sans être vicieuse...

réduction de $\bar{A}$ Il s'agit de prendre la dimension du système inerte inclus dans $\bar{A}$. Il faudra considérer ses développements de Schür en éléments propres ou singuliers et effectuer une restriction sur les valeurs les plus faibles, déterminée par une référence physique d'un temps caractéristique. Cette approche devrait s'enrichir d'une méthode dite de Hankel qui vise à réduire un modèle en équilibrant à la fois une structure optimale et une excitabilité maximale à un bruit blanc. C'est par cet aspect que l'on rejoint une problématique connexe de l'unsecte, celle de l'expérimentation et de Kalman.

résolution de la conjecture Quand on aura pu avancer sur ces premiers développements, on aura les outils pour décomposer un système complet en ses parties réduites lente-rapide. Il restera encore à justifier l'assimilation de l'effet des rapides au lent par un filtre non-L de bruit.

Il reste donc du pain sur la planche, et j'aimerais savoir qui cela intéresse-t-il d'y travailler.

Note : avec la décomposition $A(x) = \bar{A} + N(x)$, on pourra se baser sur l'ensemble des points de capiton $x(t)/A(x) \simeq \bar{A}$, car on a alors $\lim_{x \to 0} \frac{\|N(x)\|}{\|x\|} = 0$ (les points de capiton existent du fait du théorème de repliement ?).

6.2 Lenteur, inertie, et autonomie $[\phi, \theta]$ *je suis lente et pourtant j'évolue*

[1] Nous sommes ici dans la couche physique ; il semble donc pertinent d'analyser des exemples pour concrétiser les questions auquelles renvoient ces notions, au départ vernaculaires, et aboutir à des définitions formelles.

à la pêche :

- un système inerte est difficilement perturbé ;
- un gyroscope – comportant un élément en rotation rapide – est un système inerte (et même dit inertiel) ;
- une grande constante de temps (λ_{min}) peut-elle être liée à un V.P tournant rapidement ?
- la notion d'inertie est-elle plutôt liée à l'énergie interne du système ? (ce qui n'est pas le cas pour la définition du moment d'inertie, mais on parle "d'inertie chimique") ;
- rapport avec le tenseur d'inertie ou le moment d'inertie (et système inertiel)
- centre d'inertie, moment d'inertie, force d'inertie d'entraînement, de Coriolis ;
- "le temps ne mesure que le mouvement et l'inertie" Aristote, (*Physique* XIX, 11) ;
- l'inertie des investissements, du capital. Pat me dis que ça s'estime par un coût ;
- optimisation inertielle (mesure *a priori* des échelles de temps ou d'inertie) ;
- théorème d'inertie de Sylvester (?) qui semble plutôt renvoyer à une conservation ;
- lien avec la notion de passivité ? (réponse à une commande) ;
- est-on d'accord qu'un système statique est sans dynamique ?

[1] *le Cantique des Quantiques ?*

– je suis plus sûr de rien : vous voyez une inductance comme inerte ou pas ? oui ? (inertie = stock interne)

Bref, la notion d'inertie a été surtout invoquée pour la masse et a mené aux théories de champ la reliant à la courbure locale, si bien qu'on peut se demander si cette notion ne nous est pas facilement réappropriable ... non ? tant pis.

Il ne semble ainsi pas du tout évident que cette recherche soir débile, à savoir définir une notion d'inertie dans le cadre des systèmes dynamiques ; surprenant, non ? La généralisation n'exclue pas les analogies avec les approches mécaniques. En gros, on pourrait être amené à la limite à traiter le SLTC comme l'espace relativiste ... quand je vous le disais, qu'il y avait du trièdre de Frenet à y mettre, je me sentais bien traité de débile, eh bien quelqu'un l'a fait (thèse de Jean-Marc Ginoux).

(ça intéresse encore quelqu'un ?)

exercice de lecture conseillée pour élever le débat (avec une bonne surprise), extrait de présentation sur la théorie de la viabilité (lastre.asso.fr/aubin/Viabilite-Synthese.doc).

C'est moi qui colorie et qui bas-de-page.

d. Lois de Régulation des Systèmes

Dans les cas cybernétiques et cybernétiques/incertains, on cherche non seulement les états initiaux, mais également les correspondances de régulation associant à chaque état initial l'ensemble des commandes ou des régulons qui gouvernent les évolutions devant satisfaire les propriétés de viabilité et/ou de capturabilité prescrites (voir figure donnant la correspondance de régulation de la direction d'un robot devant atteindre en temps minimal la cible rouge en évitant les obstacles noirs). Ces correspondances sont étudiées théoriquement et calculées numériquement, et sont l'objet de théorèmes mathématiques difficiles.

La taille de ces correspondances de régulation fournit une mesure de la robustesse ou de la résilience de la viabilité d'un environnement par rapport au système évolutionnaire auquel il est confronté. Les dérivées de cette correspondance (c'est à propos de ce problème qu'a été inventé en 1980 le calcul différentiel des correspondances qui a par la suite joué un rôle important dans le développement de l'analyse multivoque) révèlent les propriétés de changement de structure du système évolutionnaire sous-jacentes nécessaires au maintien de la viabilité. Tant que la vitesse nulle appartient à la dérivée de la correspondance de régulation, on peut maintenir constante la variable cybernétique (commande ou régulon), tandis que le changement de cette variable s'impose dès que la vitesse nulle cesse d'y appartenir. C'est alors qu' apparaît ce que l'on pourrait appeler une « crise de viabilité » et « l'émergence de nouveaux régulons », crise d'autant plus importante que les vitesses des variables cybernétiques sont élevées. Il se peut même qu'elles deviennent infinies, auquel cas on doit recourir à des impulsions sur les variables cybernétiques et, dans des cas extrêmes, sur les variables d'état elles-mêmes. Il s'agit alors de mesurer a priori à chaque instant initial la plus grande des vitesses futures des commandes ou des régulons nécessaires au maintien de la viabilité par une fonction d'inertie qui procure ainsi une mesure des « coûts de transition » (voir figure 1). Leur étude permet également de savoir où et quand apparaîtront les crises de viabilité.

d. Le Principe d'Inertie

Le principe d'inertie énonce que les régulons n'évoluent que lorsque la viabilité est en jeu. Ces régulons, génotypes, prix, codes culturels et concepts dans les

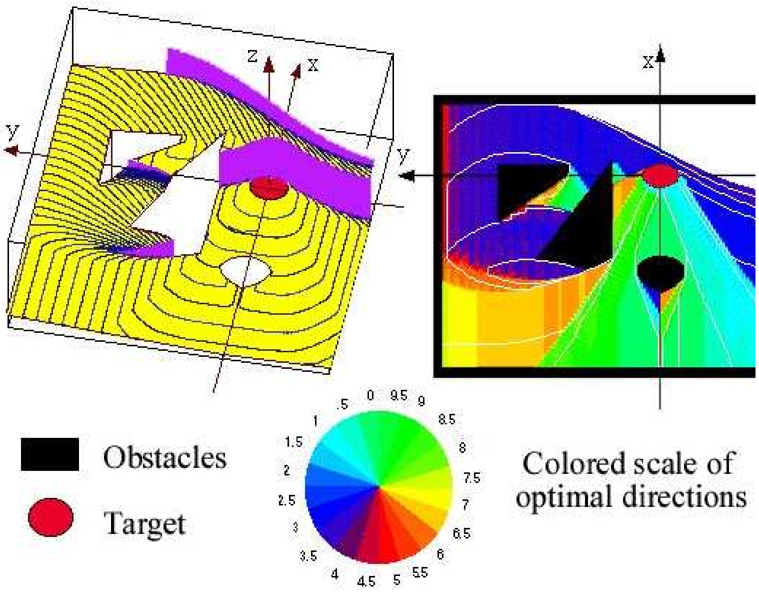


FIG. 1 – la fonction d'inertie procure une mesure des coûts de transition.

exemples cités, laissés à eux-mêmes, auront donc tendance à demeurer constants durant certaines périodes, ce qui n'interdit pas aux états du système d'évoluer. Le principe d'inertie explique ainsi le phénomène d'équilibre intermittent (punctuated equilibrium) introduit en 1972 par Nils Eldredge et Stephen J. Gould en paléontologie pour décrire des discontinuités des témoignages de l'évolution des espèces, ou l'apparition de crises historiques selon Jean-Baptiste Duroselle.

Cela ne suffit pas encore à réduire complètement l'incertitude : il faut découvrir des mécanismes qui obéissent à ce principe d'inertie. Le plus simple est celui qui consiste à choisir parmi tous les régulons viables celui qui a la plus petite vitesse, le plus paresseux. Les évolutions viables correspondantes sont qualifiées de lourdes (voir figure 2, le régulon est en rouge et l'évolution lourde en bleu).

FIG. 2 – régulon en rouge et évolution lourde en bleu.

On introduit alors le concept de niche de viabilité d'un régulon, qui est l'ensemble des états de départ d'évolutions viables régulées par ce régulon lorsqu'il est maintenu constant. La niche de viabilité d'un régulon peut être vide, auquel cas le régulon devra être changé tôt ou tard afin que l'état du système respecte les contraintes de *viabilité. Partant de la niche de viabilité d'un régulon, au moins une évolution régie par le système évolutionnaire peut évoluer dans cette niche en conservant le régulon initial. En respectant le principe d'inertie, les régulons ne se mettent en mouvement que lorsqu'une « crise de viabilité » survient, et ce, jusqu'à ce que la viabilité soit rétablie. Lorsque l'évolution lourde d'un état le conduit dans la niche de viabilité d'un régulon, alors ce dernier devient constant et l'état demeure à jamais dans sa niche, qui « accroche » (lock-in en anglais) cette évolution.*

En sélectionnant des évolutions viables lourdes, par exemple, on fait donc émerger de la confrontation de la dynamique d'un système et des contraintes qui lui sont imposées un mécanisme d'évolution opportuniste, conservateur et paresseux.

e. Échelles d'Inertie

C'est à ce sujet que se posent les problèmes d'échelles d'inertie et de temps. On cherche d'abord les situations où les régulons demeurent constants, puis celles où les régulons sont linéaires par rapport au temps, puis quadratiques, puis polynomiaux de degrés successifs. Chaque situation correspond au cas où les dérivées des régulons sont nulles jusqu'à un certain ordre, condition qui décrit une échelle donnée d' inertie, et donc une hiérarchisation en zones d'inertie décroissante. La recherche de ces zones répond à une autre préoccupation apparue en automatique sous le label de « quantisation ». Comme la recherche des rétroactions est difficile et nécessite de nombreux calculs, il s'agit d'isoler les cas où commandes et régulons dépendent d'un (petit) nombre fini de paramètres (un dans le cas des régulons constants, deux dans celui des régulons linéaires, trois dans le cas de régulons quadratiques, etc.) En effet, la recherche des rétroactions (servomécanismes, feedbacks) statiques et dynamiques revient à rechercher les sélections (univoques) de ces correspondances de régulation et de leurs dérivées respectivement. Cela peut exiger de coûteux calculs.

Ce n'est pas ce qui se fait actuellement dans la pratique en automatique. En effet, dans le cadre des systèmes linéaires dans lequel se sont développés les concepts et les techniques de l'automatique et de la théorie des systèmes, ces rétroactions sont simples et de nombreuses méthodes ont été mises au point depuis les débuts de l'automatique pour en donner des formules analytiques. Il est ensuite implicitement et parfois explicitement admis que ces mêmes rétroactions peuvent être valides dans le cas des systèmes qui ne sont pas trop « non-linéaires », sans véritables garanties mathématiques[1]. On peut concilier l'usage conjoint de ces deux points de vue, en cherchant les noyaux de viabilité des rétroactions classiques et déjà utilisées dans l'industrie, et ensuite, les utiliser comme cibles et chercher alors des feedbacks non linéaires qui y conduisent en temps fini pour utiliser ensuite les servomécanismes classiques. Cette stratégie est aussi utilisée dans le cadre des systèmes issus du vivant et des systèmes cognitifs, lorsque certaines tâches ou programmes sont appris et que le maintien de la viabilité ou l'accomplissement de tâches se décomposent en une suite « planifiée » de tâches routinières successives. Ce thème de recherche est récemment apparu dans diverses théories de l'apprentissage.

Texte édifiant malgré son encombrement par le vocabulaire de la viabilité, qui renvoie à une possibilité de combiner inertie et contrôlabilité, inertie et gains de rétroaction ! ce qu'on peut en extraire :

– maintien de la viabilité par une fonction d'inertie : l'inertie est liée au maintien spontané d'une propriété ou d'un état — sauf en cas de crise ;

[1]merci pour la publi PRL.

– on peut associer un coût (de rattrapage) à cette notion d'inertie ; (mais je me demande s'il n'y pas là confusion avec stabilité)
– Inertie ci-dessus associée à un coût de commande nécessaire au maintien de quelque chose (piste) ; avec hiérarchie possible par rapport à la forme de la loi de commande (polynômes de Legendre) ; et ce maintien peut aussi être recherché par des rétroactions (actives). l'échelle de l'inertie correspond à une dérivée s'annullant jusqu'à un certain ordre, lié au nombre de paramètres nécessaires à la régulation ;

cela suffit-il à se mettre au travail ?

6.3 réduction de modèle $[\theta, \mu]$

On va ici exposer la technique formelle de décomposition dite de Hankel en tant que référence à nos propres développements. Ce qui permettra aussi d'expliciter des liens avec la conjecture de G&M.

Il vaut tout d'abord mieux remarquer la différence de problématique avec la décomposition lent-rapide. Ici, il s'agit de réduire l'ordre d'un modèle en conservant au mieux ses caractéristiques dynamiques (on est donc en amont du lent-rapide). Celles-ci sont censées ne pas être liées au contexte (de forçage), et on considérera plutôt un forçage par bruit blanc. Alors, la structure dynamique pourra être caractérisée par une moyenne d'ensemble de la matrice de var-covar $P = \overline{\mid x >< x \mid}$ — pour rester dans le contexte SLTC. L'obtention par le modèle réduit de la plus grande partie de $\overline{< x \mid x >}$ constitue la part quantitative de cette caractéristique.

On considère ainsi le SLTC suivant :

$$\left\{ \begin{array}{l} \partial_t x = Ax + Bw(t) \\ y = C^\dagger x \end{array} \right. \tag{45}$$

où w est un bruit blanc ($Esp \mid w(t) >< w(\tau) \mid = I\delta(t-\tau)$. On a d'emblée complété le modèle d'état par une observable y. La première matrice considérée est le **grammien de contrôlabilité** (cf Effluents $\mathcal{W}_c$) :

$$P = \lim_{t\to\infty} \int_0^t \Phi(t,\tau)B(\tau)B^\dagger(\tau)\Phi^\dagger(t,\tau)\, d\tau \tag{46}$$

avec, ainsi, la supposition d'existence de cette limite, condition équivalente au critère de stabilité asymptotique dont en tire la condition de Lyapunov ($AP + PA^\dagger = -Q$) pour les systèmes autonomes[1]. Cette matrice donne la structure (les EOFs) de réponse du système au bruit non biaisé w. Ayant caractérisé le système du point de vue de sa réponse (en état, ou réponse "interne"), on veut contrebalancer en prenant également en compte le point de vue de l'observation du système via y. A ce point de vue correspond le **grammien d'observabilité**

$$Q = \lim_{t\to\infty} \int_0^t \Phi^\dagger(t,\tau)C(\tau)C^\dagger(\tau)\Phi(t,\tau)\, d\tau \tag{47}$$

[1]pour un système asymptotiquement stable (intrinsèquement, c'est-à-dire sans commande u), on a ce résultat que pour toute matrice sym.déf.pos. Q, $\exists P$ tel que $A^\dagger P + PA = -Q$.

Il suffit de prendre $P = \displaystyle\int_0^\infty e^{A^\dagger t} Q e^{At} dt$. On remarque que $A^\dagger P + PA = d_t P$ sous l'intégrale et $\lim_{t\to\infty} e^{A^\dagger t} Q e^{At} = 0$, dès que $\Re(\lambda) \in \Lambda(A) < 0$ (stabilité stricte).

L'échange entre A et $A^\dagger$ dans P montre qu'on a tout aussi bien une (autre) P telle que : $AP + PA^\dagger = -Q$.

Équilibrer une réduction de ce modèle signifie appliquer un critère équivalent selon les deux points de vue. Formellement, la structure particulière du modèle en sa version autonome autorise l'existence d'une base commune dans laquelle les deux grammiens conjugués sont égaux. Cette base donne lieu à la décomposition de Hankel.

Soit T cette transformation **interne**, c'est-à-dire un changement d'échelle des variables d'état. On **ne touche pas à la mesure**, qui est ce qu'elle est avec ses contraintes techniques conduisant à la matrice $C^\dagger$ qui fixe l'échelle des mesures. Le système originel devient ainsi, avec $x = Tz$ définissant l'état renormalisé z :

$$\left\{ \begin{array}{ll} \partial_t T z & = ATz + Bw(t) \\ y & = C^\dagger T z \end{array} \right.$$

soit

$$\left\{ \begin{array}{ll} \partial_t z & = T^{-1}ATz + T^{-1}Bw(t) = \tilde{A}z + \tilde{B}w(t) \\ y & = \tilde{C}^\dagger\, z \end{array} \right.$$

Dans le cas où A est c^{ste} où $\Phi(t,0) = e^{At}$, on en déduit les formes transormées de P et Q :

$$\left\{ \begin{array}{l} \tilde{P} = T^{-1}PT^{-\dagger} \\ \tilde{Q} = T^\dagger\, P\, T \end{array} \right.$$

Il reste à choisir T tel que $\tilde{P} = \tilde{Q} = [\sigma^2]_{diag}$. Si c'est possible, alors on doit avoir $\tilde{Q}\tilde{P} = T^{-1}QPT$, d'où $QP = T[\sigma^2]_{diag}T^{-1}$. Cette forme est la décomposition en EP ou SVD de QP, matrice produit de matrices quadratiques. Les σ_i (réels) sont appelés valeurs singulières de Hankel (v.s.H).

Noter qu'il **ne s'agit pas** des éléments propres des grammiens, car on a

$$\left\{ \begin{array}{l} P = T\,[\sigma]\,T^\dagger \\ Q = T^{-\dagger}[\sigma]T^{-1} \end{array} \right.$$

Dans le cas où on remplace le bruit w par une commande $u(t)$, cas traité par les auteurs de la présentation adoptée ici[1], les v.s.H ordonnées fournissent une borne supérieure de l'erreur de troncation à l'ordre r :

$$\|y(t) - y_r(t)\| \leq 2\Sigma_{r+1}^n \sigma_i \|u(t)\| \tag{48}$$

La réduction est alors obtenue en ne retenant que les r premières lignes et colonnes de $\tilde{A}$. Ceci est justifié par l'équation de Lyapunov liée à P.

On a d'abord $AP = T\tilde{A}T^{-1}T\tilde{P}T^\dagger = T\tilde{A}[\sigma]T^\dagger$, et $PA^\dagger = T\tilde{P}T^\dagger T^{-\dagger}\tilde{A}T^\dagger = T[\sigma]\tilde{A}^\dagger T^\dagger$;

Alors, si le système vérifie la relation de Lyapunov $AP + PA^\dagger = -I$, on a $\tilde{A}[\sigma] + [\sigma]\tilde{A} = -SS^\dagger$, avec $S = T^{-1}$. On peut donc réduire le système en gardant :

$$\tilde{A}_r[\sigma]_r + [\sigma]_r\tilde{A}_r = -S_rS_r^\dagger \tag{49}$$

où on a gardé les r premières lignes de S. Le modèle réduit se conjugue alors par $\partial_t z_r = \tilde{A}_r z_r + S_r B u$. Les mêmes composantes de u apparaissent donc dans une combinaison linéaire réduite. Dans le cas du bruit blanc non biaisé ($B = F$ unitaire), la source de bruit se met à prendre des couleurs dans la base de Hankel — ce qui

[1] Willcox Peraire, AIAA Journal 2002 (40-11).

nous rapproche un peu de G&M, sauf qu'en restant en linéaire, on aura du mal à biaiser ... ce qui démontre une fois encore la pauvreté sémantique du linéaire.

De quelques équivalences. On peut formuler l'équilibrage e/s en cherchant le maximum du **quotient de Hankel** $H(x_0) = \max_u \frac{\int_0^\infty < x(t) \mid x(t) >}{\int_\infty^0 < u(t) \mid u(t) >}$, dont on vérifie que c'est $H(x_0) = \frac{< x_0 \mid Q \mid x_0 >}{< x_0 \mid P^{-1} \mid x_0 >}$.

L'équilibrage met ainsi en branle un système commandé à $t < 0$ par la loi de commande de coût minimum nécessaire pour obtenir une variation d'état x_0 à $t = 0$ (critère de sensibilité au forçage), suivi du système libre observé aux temps positifs en son intégrale quadratique de réponse (critère de sensibilité aux C.I.). On peut voir ce système comme résultant d'un **opérateur linéaire de Hankel** :

$$\mathcal{H}[u](t) = x(t) = e^{At} \int_{-\infty}^{0} e^{-A\tau} \tilde{u}(\tau) d\tau \tag{50}$$

En considérant la SVD de l'opérateur $H = U\sigma V^\dagger$ on voit assez vite que les v.s.H sont aussi celles des extremum du quotient de Hankel et stationnarisant l'opérateur (aux carrés près).

Que dire du cas non-autonome ?

La définition des deux grammiens (46 et 47) questionne la clé des simplifications pour le système autonome reposant sur la relation de Lyapunov :

$$\begin{cases} AP + PA^\dagger = -BB^\dagger \\ A^\dagger Q + QA = -CC^\dagger \end{cases}$$

On peut étudier la relation $AP + PA^\dagger = d_t P(t, \tau)$ en remarquant que les matrices A et B sont fonctions de τ, pas de t comme pour le cas autonome où seul intervient en réalité $(t - \tau)$ paramètre $(\to t)$ de l'intégration[1]. Il reste alors $P = \lim_{t\to\infty} \int_0^t d_t \Phi B B^\dagger \Phi^\dagger(t, \tau) d\tau$ dont il n'est pas aussi évident de prendre les hypothèses de stabilité asymptotique.

On a aussi une autre relation :

$$\begin{cases} d_t P(t) = A(t)P(t) + P(t)A^\dagger(t) + BB^\dagger(t) \\ d_t Q(t) = A^\dagger(t)Q(t) + Q(t)A(t) + CC^\dagger(t) \end{cases}$$

portant cette fois sur l'intégrale de P (resp. Q) de 0 à t. On peut alors faire la remarque que ce terme de droite n'est autre que la source faisant évoluer la matrice de var-covar P d'un modèle (cf Propagateur equ. (186)) sollicité par un bruit blanc (Bw). En gros dans le cas non-autonome, on devra s'assurer lors de la réduction du modèle que le modèle réduit n'est pas moins stable que le modèle complet en considérant ces termes source et leur limite à l'infini.

Enfin, cette dualité entre P et Q[2] suggère que l'on puisse ajouter au système le calcul d'une avance de Q pour déterminer $< y \mid y >$ par une expression du même type que celle de dP (equ. (186) de Propagateurs) :

$$dQ = [I - \tau A^\dagger]^{-1} \left[A^\dagger(t)Q(t) + Q(t)A(t) + C^\dagger RC \right] [I - \tau A]^{-1} \delta t \tag{51}$$

[1] on a appelé $P(t, \tau)$ l'intégrande.

[2] ben non, ça marche pas pour Q, par le fait que A ne commute avec Φ que dans le cas autonome !

ou pas très éloignée de ça, avec W la matrice de var-covar des incertitudes de mesure. Contrairement à la fausse relation (51), celle-ci marche puisqu'on assimile localement le propagateur à $\exp(A\tau)$. Bien sûr au vu de cette forme, on doit pouvoir avoir le droit d'affirmer que cette Q n'est autre que la P des variables adjointes aux états, et on aura à y réfléchir pour les implémentation dans Mini_ker.

Une hiérarchie de réduction

En résumé, on peut établir une série de décompositions menant à une hiérarchie de réductions de modèle possible :

- *modale* - $\Phi = \sum_i e^{\lambda_i t} | e_i >< e_i |$; on tronque â partir de $\Re\lambda_i$,
- *modale*
- *modale*
- *modale*

un exercice d'application devrait permettre d'apprécier la différence des résultats entre ces diverses formes, et en particulier en fonction de la normalité de la matrice d'avance A.

7 lent / rapide, encore

Au hasard des lectures, on découvre une école se référent à la technique des projecteurs introduite par R. Zwanzig et S. Nakajima dans les années 60, exploité par des J. Seke *et al* plus tardivement (années 80), puis enfin relié aux stochastiques plus récemment (cf P. Espanol er F. Vazquez 20020. La technique permet de poser un cadre général de séparation formelle, ici appliquée au lent / rapide, mais qui — me semble-t-il — devrait avoir un lien avec l'approche Materiochka qui exploite une décomposition (spatiale) itérée à la Reynolds. L'idée de base consiste en l'application d'un projecteur constant $\mathcal{P}$ sélectionnant dans l'espace des phases (des états) une partie dîte pertinente — de dimension très réduite ou directe ou quasi déterministe — et son complémentaire $\mathcal{Q} = I - \mathcal{P}$ donnant la partie bruitée ou non-pertinente. De manière plus générale, l'approche peut sélectionner des "niveaux de description mésoscopiques"

Le problème est que cette méthode est appliquée dans des contextes assez divers, des modèles de colloïdes genre Fokker-Planck aux équations de Liouville ou Schrôdinger, et je tente d'aller lentement sur une forme la plus générale avant de considérer les applications traitées.

Une première voie consiste a chercher une "équation maîtresse généralisée[1]"

Supposons défini tel projecteur ; il est appliqué à l'équation d'évolution — ici hors TEF :

$$d_t x(t) = A(t)x \quad , \quad x(0) = x_0; \tag{52}$$

avec d'abord une partie pertinente

$$d_t \mathcal{P}x(t) = \mathcal{P}A(t)x = \mathcal{P}A\mathcal{P}\ \mathcal{P}x + \mathcal{P}A\mathcal{Q}\ \mathcal{Q}x \tag{53}$$

où après produit par $\mathcal{P}$ à droite — qui commute avec d_t — on a utilisé : $\mathcal{P} + \mathcal{Q} = I$, $\mathcal{P}^2 = \mathcal{P}$, $\mathcal{Q}^2 = \mathcal{P}$. On peut former la même équation pour le versant complémentaire de l'espace des états :

$$d_t \mathcal{Q}x(t) = \mathcal{Q}A\mathcal{P}\ \mathcal{P}x + \mathcal{Q}A\mathcal{Q}\ \mathcal{Q}x \tag{54}$$

Dès lors, la sémantique recherchée nous conduit renommer par exemple $\mathcal{P}x$ en x et $\mathcal{Q}x \to y$, ce qui donne un système partitionné :

$$\begin{cases} d_t x = \mathcal{P}A\mathcal{P}\ x + \mathcal{P}A\mathcal{Q}\ y \\ d_t y = \mathcal{Q}A\mathcal{P}\ x + \mathcal{Q}A\mathcal{Q}\ y \end{cases}$$

avec $x \in \Re^n$ et $y \in \Re^m$, $n + m$ donnant la dimension totale de l'espace des états.

Il est alors temps d'introduire une différence de traitement entre les deux partitions ; y étant "non-pertinente" doit pouvoir être traitée globalement ou statistiquement comme fluctuations (sous-maille par exemple). On résoud formellement $y(t)$ sous des conditions sur lesquelles nous reviendrons, mais on va ici supposer $A(t) = C^{\text{ste}}$, et raisonner sur les états avant de passser a l'étape opérateurs.

$$y(t) = \exp(\mathcal{Q}A\mathcal{Q}\ t)\ y(0) + \int_0^t \exp(\mathcal{Q}A\mathcal{Q}(t-\tau))\ \mathcal{Q}A\mathcal{P}\ x(\tau)d\tau \tag{55}$$

[1] école de Bruxelles. cf Brandtlow et Conveney, et Penrose.

L'approche clalssique consiste à reporter cette expression dans la partie pertinente pour obtenir une "équation maîtresse généralisée" :

$$d_t x(t) = \mathcal{P}A\mathcal{P}x + \mathcal{P}A\mathcal{Q}\exp(t\mathcal{Q}A\mathcal{Q})y_0 + \mathcal{P}A\mathcal{Q}\int_0^t \exp(\mathcal{Q}A\mathcal{Q}(t-\tau))\ \mathcal{Q}A\mathcal{P}\ x(\tau)d\tau \tag{56}$$

équation dîte maîtresse.

attention à la marche , car on passe aux opérateurs. Soit la dynamique microscopique d'un système donnée par un groupe d'opérateurs $\mathcal{U}_t$: $x'(t) = \mathcal{U}_t x'(0)$ dans l'espace complet. On note $\mathcal{L}$ le générateur infinitésimal du groupe, avec ainsi :

$$d_t \mathcal{U}_t x'_0 = \mathcal{L}\mathcal{U}_t x'_0\ ,\ x'_0 \in \mathcal{D}(\mathcal{L}); \tag{57}$$

où $\mathcal{D}(\mathcal{L})$ est le domaine de $\mathcal{L}$. Alors, $x'(t) = \mathcal{U}_t x'(0)$ est une solution continument différentiable de l'équation

$$d_t x'(t) = \mathcal{L}x'(t); \tag{58}$$

en réalité, je m'avance, parce que ce qui précède est établi pour le groupe des opérateurs unitaires $\mathcal{U}_t$ et $\mathcal{L}_t$ anti-adjoint ; il s'agit d'obtenir un groupe de $\mathcal{U}_t$ "strongly continuous[1]". C'est sûr que l'application d'opérateurs conservant la mesure facilite les choses, mais outre-passons *à la physicienne.*

On réapplique les projections, ce qui revient à remplacer toute occurence de $A(t)$ par $\mathcal{L}_t$. On arrive alors formellement à :

$$d_t x = \mathcal{P}\mathcal{L}\mathcal{P}x + \mathcal{P}\mathcal{L}\mathcal{Q}\exp(t\mathcal{Q}\mathcal{L}\mathcal{Q})y_0 + \int_0^t \mathcal{P}\mathcal{L}\mathcal{Q}\exp(\mathcal{Q}\mathcal{L}\mathcal{Q}(t-\tau))\ \mathcal{Q}\mathcal{L}\mathcal{P}\ x(\tau)d\tau \tag{59}$$

On peut supposer que les C.I. ont réussi à ce que $y(0) = 0$. Ici, x, la partie pertinente, est donnée par la projection $\mathcal{P}x'(t) = \mathcal{P}\mathcal{U}_t x'_0 = \mathcal{P}\mathcal{U}_t(\mathcal{P}+\mathcal{Q})x'_0$ de l'état total x', avec $y_0 = 0$, on a $x(t) = \mathcal{P}\mathcal{U}_t\mathcal{P}x_0$. On appelle $\mathcal{V}_t = \mathcal{P}\mathcal{U}_t(\mathcal{P}$ l'opérateur d'évolution de la partie pertinente de x'.

Pour aboutir à cette fameuse **équation maîtresse généralisée**, on introduit ces notations dans (59) :

$$\begin{array}{l} d_t\mathcal{V}(t)x_0 = \mathcal{P}\mathcal{L}\mathcal{P}\mathcal{V}(t)x_0 + \int_0^t \psi(t-\tau))\ \mathcal{V}(\tau)\ x_0 d\tau \\ \textit{avec :}\ \psi(t) = \mathcal{P}\mathcal{L}\mathcal{Q}\exp(t\mathcal{Q}\mathcal{L}\mathcal{Q})\ \mathcal{Q}\mathcal{L}\mathcal{P} \end{array} \tag{60}$$

qui est appellé "opérateur de collisions" par l'école de Bruxelles. Cette équation est alors une intEgro-différentielle fermée pour la part pertinente de l'évolution de l'état total.

Nous qui sommes du TEF, on écrirais plus volontiers :

$$\begin{cases} d_t\mathcal{V}(t)x_0 = \mathcal{P}\mathcal{L}\mathcal{P}\mathcal{V}(t)x_0 + \mathcal{P}\mathcal{L}\mathcal{Q}\ \overline{y}(t) \\ \overline{y}(t) = \int_0^t \psi(t-\tau))\ \mathcal{V}(\tau)\ x_0 d\tau \end{cases} \tag{61}$$

mais avec : $\psi(t) = \exp(t\mathcal{Q}\mathcal{L}\mathcal{Q})\ \mathcal{Q}\mathcal{L}\mathcal{P}$, c'est-à-dire que la partie fluctuante a été (g)lissée dans un transfert.

On imagine bien que cette grande généralité (en réalité restreinte quand même aux opérateurs unitaires susdîts) soit la source de recherches sur des systèmes

[1] je laisse en Anglais pour que Guillaume nous donne la juste expression française.

traitables pratiquement. Une analyse formelle en transformées de Laplace des opérateurs, pour analyser le comportement asymptotique de $\mathcal{V}(t)$ conduit l'école de Bruxelles à la notion d'opérateur de collision à décroissance exponentielle :

$$\exists K > 0,\ b > 0,\ \textit{tel que}\ ||\psi(t)|| \leq K \exp^{-bt} \quad (t \in \Re^+) \tag{62}$$

avec $||.||$ une norme sur les opérateurs. On doit alors a Coveney et Penrose ce résultat que dans ces conditions, et si $\mathcal{P}$ est de rang fini, on peut décomposer cette évolution en $\mathcal{V}(t) = \mathcal{W}(t) + \mathcal{W}'(t)$, avec les propriétés :

1. $\exists K' > 0,\ \textit{tel que}\ ||\mathcal{W}'(t)|| \leq K' \exp^{-\frac{bt}{2}} \quad (t \in \Re^+)$;
2. $\mathcal{W}(t) = \exp(t\mathcal{P}\Gamma\mathcal{P})$, oú $\mathcal{P}\Gamma\mathcal{P}$ est un opérateur de rang fini de spectre compris dans $\{z \mid -\frac{b}{2} \leq \Re(z) \leq 0\}$;

Ainsi, on a d'une part lissé la partie non-pertinente poour la faire disparaître, et montré une condition suffisante pour obtenir un comportement asymptotique stable en présence des fluctuations.

En ce qui nous concerne plus sobrement, tout en gardant un œil sur les diverses élucubrations tempestives de nos amis stati-stochas — dont nos amis Belges — on pourra en rester à l'interprétation SLTCive et interroger la forme :

$$\left\{ \begin{array}{l} d_t x(t) = \mathcal{P}A(t)\mathcal{P}x(t) + \mathcal{P}A(t)\mathcal{Q}\,\overline{y}(t) \\ \overline{y}(t) = \int_0^t \exp(\mathcal{Q}A(?)\mathcal{Q}(t-\tau))\ \mathcal{Q}A(\tau)\mathcal{P}\ x(\tau)d\tau \end{array} \right. \tag{63}$$

On sent bien que dans cette forme on a fait l'essentiel consistant à transformer les fluctuations en un transfert calmé par une prise de moyenne pondérée, et une éventuelle décroissance exponentielle, mais il reste quelques points d'interrogation[1]. En tout cas, on voit où ce situe tout le tripatouillage à faire pour construire un système physique "lent" ayant du sens...

une application charnière , en prenant un exercice de Majda, détourné pour en faire un traitement intégral. On considère le système suivant :

$$\left\{ \begin{array}{l} d_t x = a\,yz \\ d_t y = b\,xz \\ dz = c\,xy\,dt - \gamma z\,dt + \sigma\,dW \end{array} \right. \tag{64}$$

où on a écrit la dernière équation à la Itô. On suppose $a + b + c = 0$ pour en faire "la triade" stochastique nonlinéaire (réputée chez les stochamologues). Par rapport à ce qui précède, l'analyse de pertinence a donc déjà été effectuée, il reste à passer à l'échelle temporelle supérieure (mesoscopique). On reviendra ensuite à une comparaison plus ponctuelle avec la méthode générale.

La troisième équation, à part un forçage par les lents, a la forme d'une équation de Langevin. On souhaite transformer ce système en une forme où les lents verraient une moyenne du rapide, le rapide étant forcé par les lents. Les lents constituant une source pour $d_t z$, la solution formelle de $z(t)$ est donnée par :

$$\begin{array}{l} z(t) = c\int_0^t \exp(-\gamma(t-\tau))x(\tau)y(\tau)\,d\tau + g(t) \\ \textit{avec}\ g(t) = \sigma\int_0^t \exp(-\gamma(t-\tau))\,dW(\tau) \end{array} \tag{65}$$

[1] tout ceci, je dois l'avouer, pour ne pas en passer par le fâcheux $t \to \epsilon\tau$ de Majda, qui me gonfle à un point suspect : "small is not necessarily beautiful". Mon plus grand succès serait de montrer l'équivalence des deux approches ... mais brutalement me revient en tête le papier de Michael sur la moyenne SANS petit paramètre, par utilisation de la transformation de Lie ¿

et ainsi, après remplacement dans le système des lents, les fluctuations rapides seront lissées par une prise de moyenne d'ensemble avec une moyenne pondérée des intégrations, sur lesquelles les hypothèses physiques doivent être faîtes. L'analyse de la source de bruit lissé $g(t)$ se fait classiquement en étudiant l'espérance de la fonction de corrélation :

$$\rho(t,t') = \overline{g(t)g(t')} = \sigma^2 \int_0^t \int_0^{t'} \exp\{-\gamma(t-\tau)\} \exp\{-\gamma\{t'-\tau'\}\} \overline{dW(\tau)dW(\tau')} \tag{66}$$

le processus de Wiener est caractérisé par $\overline{dW(t)dW(s)} = \delta(t-s)\,dt$, ce qui donne, avec $t' \geq t$:

$$\begin{aligned} \rho(t,t') &= \sigma^2 \int_0^t \exp\{-\gamma(t'+t-2\tau)\}d\tau \\ &= \frac{\sigma^2}{2\gamma}[\exp(-\gamma(t'-t)) - \exp(-\gamma(t'+t))] \end{aligned} \tag{67}$$

Bon, mais ce terme en $t'+t$ est troublant, et il semble que l'introduction d'une stationnarité de $\rho(t,t)$ conduise à prendre une forte décroissance exponentielle $\gamma \to \frac{\gamma}{\epsilon}$, ce qui implique, pour garder une corrélation finie, son pendant $\sigma^2 \to \frac{\sigma^2}{\epsilon}$, ou qui peut m'en dire mieux ? On doit en effet garder en tête qu'on n'a pas à étudier de systèmes stochastiques arbitrairement idiots mais au contraire se donner des règles pour en construire certains produisant du sens physique...

Alors, doit-conclure à un échec, dans la nécessité d'avoir à introduire un petit paramètre ? Ce choix de "stationnarité de $\rho(t,t)$ en posant $2\gamma t \propto \epsilon$ peut aussi s'interprèter par $t \to \frac{t}{\epsilon}$, et on a "considéré des temps longs" par la variable $\epsilon\, s = t$, justifiant cette méthode. Au moins y avons-nous été contraints par une sémantique physique ?

Examinons les conséquences de cette hypothèse sur le terme intégral dans la solution de $z(t)$, en exhibant la lenteur des lents :

$$\begin{aligned} z(t) &= cx(t)y(t) \int_{t-0}^t \exp\{-\frac{\gamma}{\epsilon}(t-\tau)\}\, d\tau + g(t) \\ &\overset{\epsilon \to 0}{=} \frac{\epsilon c}{\gamma} x(t)y(t) + g(t) \end{aligned} \tag{68}$$

avec la borne inférieure de l'intégrale $t-0$ qui pourra être prise à $t - \frac{\sqrt{\epsilon}}{\gamma}$. On nous fait boire jusqu'à la lie l'obligation de propager notre petit paramètre à $c \to \frac{c}{\epsilon}$ pour conserver du sens, et ainsi :

$$\begin{aligned} dz &= \frac{1}{\epsilon}[\, c\, xy\, dt - \gamma z\, dt\,] + \frac{\sigma}{\sqrt{\epsilon}} dW \\ z(t) &= \frac{c}{\gamma} x(t)y(t) + g(t) \end{aligned} \tag{69}$$

À ce stade du lissage, nous sommes en présence d'un nouveau processus aléatoire $g(t)$, dont nous pouvons caractériser le temps de corrélation en analysant, pour une fonction test $\psi(t,s)$ arbitraire :

$$\left\{ \int_0^T \int_0^T \psi(t,s)\rho(t,s)dtds \to \frac{\sigma^2}{\gamma^2} \int_0^T \psi(t,t)dt \right. \tag{70}$$

c'est ce qu'affirme Majda, sans que j'arrive à ce résultat ... et on(il) en déduit que $g(t)$ est approximativement aussi un bruit blanc avec

$$\frac{1}{\epsilon} g(t) dt \rightarrow \frac{\sigma}{\gamma} dW(t) \tag{71}$$

mais avec un temps de corrélation fini — qui tend vers zéro avec ϵ, c'est ça le "approximativement" ?.

Si on l'en croit, le système transformé est à présent :

$$\begin{cases} d_t x(t) = a\, yz \\ d_t y(t) = b\, xz \\ z(t) = \frac{c}{\gamma} xy + \frac{\sigma}{\gamma dt} \circ dW(t) \end{cases} \tag{72}$$

oú nous dit-on le signe "$\circ$" donne à interpréter ce nouveau bruit dW à la Stratonovitch. L'intervention du dt dans ce transfert marque bien l'occurence d'un batteur mésoscopique du temps, grand devant le temps de décroissance $\frac{1}{\gamma}$, en accord avec la dynamique lente.

Un question se pose à l'application du schéma temporel du TEF : comment calculer le terme source : $dz = cx\, dy + cy\, dx + ?d\{\frac{\sigma}{\gamma dt} \circ dW(\tau)\}$? **Pat, Guillaume, au(x) secours !**

Ici, les transferts du TEF s'introduisent d'une manière je trouve assez belle, sur des raisons de réponse rapide au système des lents[1].

Par rapport à l'approche générale, on a ici défini implicitement les projecteurs P et Q triviaux conservant x, y et z respectivement de l'espace des états complet. L'intégrale de lissage du bruit de la formule (63) est très simple, avec $\exp(\mathcal{Q}A(?)\mathcal{Q}(t-\tau)) \rightarrow \exp(-\gamma(t-\tau))$, et la décroissance exponentielle explicite, et $\mathcal{Q}A(\tau)\mathcal{P}x(\tau) \rightarrow cy\, dx + cx\, dy = d(cxy(\tau))^2$.

après cet effort que j'ose autoqualifier de valeureux, je craque et préfère attendre vos réponses ...

[1]ou une cellule stationnaire pour la variable t mesoscopique.

[2]attention au fait que dans l'expression (63), x est la variable dx du SLTC, sorry folks.

8 un exemple de gain

En recherchant un exemple élémentaire exhibant l'intérêt du gain circulant, on est arrivé avec stb à ce modèle de croissance d'un espèce limitée par, mettons, une maladie contagieuse par contact, formalisé par la classique logistique :

$$\partial_t \eta = r\eta(1 - \frac{\eta}{k}) \tag{73}$$

ayant comme point d'équilibre stable $\eta = k$, et dont la solution est[1]

$$\eta(t) = \frac{k\eta_0 e^{rt}}{k + \eta_0(e^{rt} - 1)} \tag{74}$$

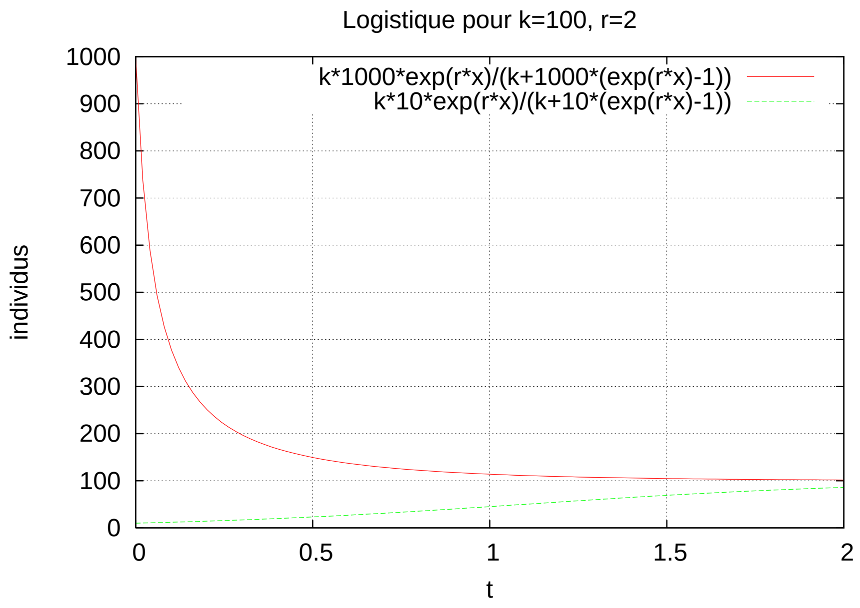


FIG. 3 – Evolution du nbre d'individus pour 1000 et 10 au démarrage.

On donne une structure de rétroaction en posant $\varphi = \eta$:

$$\left\{ \begin{array}{lcl} \partial_t \eta & = & r\eta - \frac{r}{k}\varphi^2 \\ \varphi & = & \eta \end{array} \right.$$

à présent, si la perturbation de φ, mettons négative, correspond à un nombre ϵ d'individus vaccinés : $\varphi \to \eta - \epsilon$ (le nombre des non-vaccinés), le terme non-linéaire de rencontres diminue la mortalité de $2\epsilon x$. Le nouveau point d'équilibre sur le SLTC (ϵ^2 négligé) est alors $k + 2\epsilon$: on a vacciné ϵ, on en a gagné deux[2] ! Cette interprétation est tout à fait intéressante, car là où on devait jeter des poignées de lapins pour perturber un prédateur-proie, on a simplement comptabilisé le nombre de vaccinés dès le départ qui se soustrait à la population vulnérable à chaque instant (puisque

[1] on passe par $y = \frac{1}{\eta} - \frac{1}{k}$

[2] cette interprétation innovante de la perturbation vaccinale a été inspirée à stb par la montagne Sainte-Geneviève en Juin 2007. On gagne deux parce qu'un vacciné qui rencontre un malade ne risque rien, **et** inversement.

eux, ils meurent pas en première approximation, ou alors on s'arrange pour qu'il y en ait toujours ϵ d'immunisés).

Que va nous dire l'analyse de rétroaction circulante ? Le SLTC s'écrit

$$\left\{ \begin{array}{rcl} \partial_t x & = & rx - 2\frac{r}{k}\varphi y \\ y & = & x \end{array} \right.$$

avec comme d'hab $x = \Delta\eta$, $y = \Delta\varphi$. Le propagateur correspondant est

$$\Phi(t,0) = \exp\{r(1 - 2\frac{\varphi(t)}{k})t\} \tag{75}$$

où $\varphi(t)$ est défini le long de la trajectoire de référence. On peut remarquer qu'au point d'équilibre $\Phi(t,0) = \exp(-rt)$, qui montre que ce point est bien stable — c'est-à-dire que l'analyse linéaire fournit la stabilité, bien qu'il faille dériver 73 pour conclure.

Avec ce modèle de recopie d'état ($c = 1, b = -\frac{2r}{k}\varphi$), la réponse efficace à la vaccination est simplement :

$$\rho^\epsilon(t,0) = -\int_0^t \exp\{r[t(1 - \frac{2}{k}\varphi(t)) - \tau(1 - \frac{2}{k}\varphi(\tau))]\}\frac{2r}{k}\varphi(\tau)\,\epsilon d\tau \tag{76}$$

efficacité, donc, toujours de signe contraire à ϵ.

En vaccination du point d'équilibre ($\varphi = k$), on trouve que $\rho(t,0) = -2(1 - \exp(-rt))$, qui tend vers -2 par valeurs supérieures. Qu'est-ce que ça signifie ? La variable φ comptabilise le nombre d'individus ; pour ϵ négatif, on obtient asymptotiquement 2ϵ de plus d'individus que sans vaccination[1].

[1] le “2” trouvé précédemment n'a pas de sens selon l'efficacité, il s'agit juste du déplacement du point d'équilibre stable ; mais on le retrouve ici ¿Est-ce qu'un jeune émérite écrirait l'histoire du **2** dans cette collaboration TEF-ZOOM ?

9 sur Gram-Schmidt et les petits poids

Signalée par Pat, on a une série d'algorithmes des années 74-77 par J.L. Thorton (JPL) qui concernent l'avance temporelle d'une factorisation de Cholesky dans le cadre de l'application du filtrage de Kalman discret. Si je paraphrase les notations de Thorton pour nous les rendre plus transparentes, il pose le problème suivant :

$$\begin{cases} \eta(t+\delta t) = \Phi\eta(t) + Bw \\ \mu(t) = C^\dagger\eta + n \end{cases} \tag{77}$$

équation transcrites du discret, avec w et n deux bruits indépendants de moyenne nulle et sans corrélation entre eux, avec $\overline{\mid w >< w \mid} = Q_{diag}$ notée q, $\overline{\mid n >< n \mid} = R_{diag}$ notée r et $\overline{\mid w >< n \mid} = 0$. Les états η sont dans $\Re^n$ et les probes dans $\Re^p$. Il me semble que Pat noterait $w \to dW$ pour un accroissement Brownien. Je passe la partie donnant le gain de Kalman (la partie μ ne sera donc plus évoquée), pour en arriver à la partie "time-advance".

Si $P = \overline{\mid x >< x \mid}$ est la matrice de var-covar des anomalies x, on suppose déjà effectuée la factorisation de Cholesky $P(t) = Ud\ U^\dagger$ où d est diagonale positive et U triangulaire sup. à diagonale unité. La question est de mettre à jour cette factorisation pour $P(t+\delta t)$ connaissant (77) — on peut faire le lien avec le TEF de $\Phi = e^{H\delta t}$.

Le principe est d'obtenir une première factorisation, puis de la transformer en Cholesky grâce à une orthogonalisation de G-S pondérée :

- concaténation, on pose $W = [\Phi U, B]$ et $D = diag(d, q)$ (on a donc fait avancer U) ;
- factorisation première de $P' = \Phi P \Phi^\dagger + BQ_{diag}B^\dagger = WDW^\dagger$, qui comporte bien la partie diagonale mais W n'est ni triangulaire ni unité ;
- transformation par WGS[1], en définissant la base $\mid v_i >$ de vecteurs "D-orthogonaux" avec $< u \mid v >_D = \sum_k u_k v_k D_k$, produit scalaire euclidien pondéré par la diagonale.
- D-orthogonalisation : et ainsi — si les matrices sont rangées par lignes — $< v_i \mid v_j >_D = \sum_k v_{ik} v_{jk} D_k = \delta_{i,j} || v_i^2 ||_D$ est représenté matriciellement par $VDV^\dagger = [|| v_i^2 ||_D]_d$;
- on a alors $W = U'V$ en posant U' une matrice définie par la WGS : U' est triangulaire par le procédé de G-S, rendue sup. en l'initialisant avec $\mid v_n >=\mid w_n >$;
- elle est dîte "unit triangular" par le fait de sa diagonale constituée de 1[2] ;

il en suit que $WDW^\dagger = U'VDV^\dagger U'^\dagger = U'[|| v_i^2 ||_D]_d\ U'^\dagger$, qui a la forme Cholesky, et est l'unique telle forme. G-S a ainsi permis de transformer W en un produit de matrice Cholesky par une matrice D-orthogonale ($VDV^\dagger$ =diag). L'astuce réside en une méthode qui préserve la triangularité, par G-S, tout en offrant une orthogonalité destinée à conserver la matrice diagonale (cf plus bas sur la généralisation possible d'un tel procédé de mise à jour).

démonstration On orthogonalise $\{\mid v_i >\}$ de $\{\mid w_i >\}$ à n vecteurs indépendants comme ça : (on note U au lieu de U') ;

[1] Weighted Gram-Schmidt.

[2] pas "unitaire", Pat.

- initialisation $|\, v_n >=|\, w_n >$, (à rebours pour obtenir un triangle supérieur pour U) ;
- on construit un premier vecteur $|\, v_{n-1} >$ dans $\{|\, v_n >, |\, w_{n-1} >\}$ orthogonal à $|\, v_n >$

$$|\, v_{n-1} >=|\, w_{n-1} > - \frac{< w_{n-1} \mid v_n >_D}{||v_n^2||_D} \,|\, v_n >$$

  comme G-S classique à ceci près que les produits scalaires sont pondérés — c'est-à-dire que l'on soustrait du vecteur sa (ses) projection(S) sur la base déjà formée :
- itération

$$|\, v_i >=|\, w_i > - \sum_{k=i+1}^{n} \frac{< w_i \mid v_k >_D}{||v_k^2||_D} \,|\, v_k >$$

- pour s'abstraire de la restriction sur la somme sur k, on définit la matrice triangulaire sup. U telle que $u_{i,k} = \frac{<w_i|v_k>_D}{||v_k^2||_D}$ pour $k > i+1$; om met 1 sur la diagonale pour obtenir $|\, w_i >= \sum_{k=1}^{n} u_{i,k} v_{k,j}$, soit : $W = UV$;

Pour l'histoire, cette factorisation est celle de Crout, celle de Cholesky en découlant $(LL^\dagger)$.

L'intérêt de cet algorithme réside selon l'auteur en son efficacité (éprouvée sur UNIVAC 1108 ⸮), mais pour nous, il représente en plus[1] une méthode d'avance temporelle interpellant le TEF. Par exemple, l'orthogonalisation s'effectue dans un espace de phase augmenté de B — les vecteurs ligne de W — qui ici s'applique à l'espace des bruits mais dans le TEF aux transferts **à poursuivre**[2].

Mais Thorton traite le cas particulier "d'un bruit de couleur[3]" qui cette fois, interpelle la MK2. En ce que le système a, disons, une famille de bruits blancs pondérés par une matrice $M(t)$ diagonale, insensibles aux états, que l'auteur écrit comme ça :

$$\begin{bmatrix} x \\ p \end{bmatrix} (t+\delta t) = \begin{bmatrix} \Phi_x & \Phi_{xp} \\ 0 & M \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} (t) + \begin{bmatrix} 0 \\ e \end{bmatrix}$$

Il est alors préconisé de résoudre d'abord la partie déterministe :

$$\begin{bmatrix} \overline{x} \\ \overline{p} \end{bmatrix} = \begin{bmatrix} \Phi_x & \Phi_{xp} \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} (t)$$

avant de compléter par l'avance stochastique :

$$\begin{bmatrix} x \\ p \end{bmatrix} (t+\delta t) = \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix} \begin{bmatrix} \overline{x} \\ \overline{p} \end{bmatrix} + \begin{bmatrix} 0 \\ e \end{bmatrix}$$

On va voir que les propriétés de chaque système permet de simplifier les calculs de l'avance des facteurs de P. On suppose ainsi que $P = UDU^\dagger$ avec la sous-structure états / bruits :

$$U = \begin{bmatrix} U_x & U_{xp} \\ 0 & U_p \end{bmatrix}, \quad et \quad D = [D_x, D_p]_d \,.$$

Alors, l'avance déterministe de la factorisation est simplement :

[1] On a déjà cet intérêt pour l'implémenter dans Mini_ker, non ?
[2] il y a certe le problème de l'emboîtement, mais aussi en intuiter une méthode pour faire avancer les É.P. et la SVD.
[3] il ne s'agit pas d'un bruit noir.

– $\underline{U_p}$ et D_p inchangés ;
– $\overline{\Phi_{xp}} = \Phi_x U_{xp} + \Phi_{xp} U_p$;
– $\overline{UDU}^\dagger = [\Phi_x U_x] D_x [\Phi_x U_x]^\dagger$;

on a juste un WGS à performer sur cette dernière expression, l'avance de Φ_{xp} étant une simple avance linéaire.

A présent, le système bruité est diagonal et autorise la mise à jour coefficient par coefficient. L'auteur pose les principes du traitement en considérant d'abord le problème suivant : on pose $P = \overline{UDU}^\dagger$, avec la structure :

$$\overline{U} = \begin{bmatrix} U_a & U_{ab} & U_{ac} \\ 0 & 1 & U_{bc} \\ 0 & 0 & U_c \end{bmatrix}; \quad \overline{D} = \ Diag\,[D_a, d, D_c]$$

alors en posant $\Phi = [I, m, I]_d$ et $Q = [0, q, 0]_d$ tels que $P' = \Phi P \Phi^d agger + Q$, le système bruité, on détermine U' et D' par :

– $U'_{ac} = U_{ac}$, $U'_c = U_c$ et $D'_c = D_c$;
– $d' = m^2 d + q$;
– $U'_{bc} = m U_{bc}$;
– et $U'_{ab} = \frac{d}{d'} U_{ab}$;

alors, U'_a et D'_a satisfont à l'égalité :

$$U'_a D'_a {U'_a}^\dagger = U_a D_a U_a^\dagger + (\frac{d}{d'} q) U_{ab} U_{ab}^\dagger$$

cette dernière relation est trouvée semblable à la forme traitée par l'agorithme de Agee-Turner, qui est ainsi sollicité pour une efficacité de mise à jour des coefficients. L'application au système bruité est alors :

soient n et m les dimensions de x et p.

On calcule itérativement pour $k = 1, m$

– $d'_{n+k} = m_k^2 \overline{d}_{n+k} + q_k$;
– pour $i = 1, 2, ..., n+k-1$, les vecteurs $|v>$ et$|u'>_{n+k}$:
  – $v_i = \overline{u}_{i,n+k}$;
  – $u'_{i,n+k} = m_k \frac{\overline{d}_{n+k}}{d'_{n+k}} v_i$;
– $u'_{i,n+k} = m_k \overline{u}_{n+k,j} \quad j = (n+k+1), ..., (n+m)$;

puis avec un Agee-Turner, on met à jour par :

$$ {U'}^k {D'}^k [{U'}^k]^\dagger = U^k D^k [U^k]^\dagger + c_k |\, v >< v \,|$$

où ${U'}^k$ — idem pour D —représente la partie supérieure gauche d'ordre $(n+k-1)$ des matrices U' et D', et où $c_k = \frac{\overline{d}_{n+k}}{d_{n+k}} q_k$.

Finalement encore, l'auteur fait remarquer — ce qui concerne encore MK2 — que la méthode s'étend à la présence de“ paramètres biaisés”. Si la partition est alors $[x, p, \varpi]^\dagger$, avec ces paramètres en dernier, l'algorithme s'étend en étendant la précédente boucle sur j à $n+m+$la dimension de $|\varpi>$.

applications pour MK2 ? On peut tenter de résumer les aspects de la méthode de Thorton qui puisse alimenter le principe d'emboîtement de la MK2. En gros, au lieu d'utiliser la A^{spha} comme dans la 102, faire avancer la décomposition avec une méthode ne portant que sur U et D laisse espérer un possible emboîtement.

– la toute première étape consite à faire avancer U triang. sup. unité classiquement. Soumise, donc a un calcul classique du MK2 mais on peut prendre en compte la forme spécifique de U ;
– la séparation modèle déterministe versus bruité : Pat me fait remarquer que bien qu'il soit *a priori* difficile de séparer des familles ayant une P nulle, l'occurence d'une telle possibilité pose des problèmes au filtrage. Il faut donc le permettre, et ainsi profiter des algos de séparation de Thorton.
– sur la WG-Sdation, elle ne concerne que la partie $\Phi_x U_x$ déjà calculée, et on a simplement cette trigonalisation à effectuer.

il apparaît qu'on évite, semble-t-il, la A^{spha}.

généralisation Prenant SVD par exemple, on aurait $P = UwU^\dagger$ et $P + dP = \Phi P \Phi^\dagger$, avec cette fois-ci U unitaire. Il faudrait trouver une factorisation de ΦU qui soit le produit d'une matrice unitaire par une matrice D-orthogonalisable, comme $\overline{U} V D V^\dagger \overline{U}^\dagger$, avec $\Phi U = \overline{U} V$. En fait, il faudrait effectuer une SVD de $\Phi U = XyZ^\dagger$ avec $XX^\dagger = I_n$ et $Z^\dagger D Z = [||z_i||^2_D]_d$ afin que $P = \Phi U D [\Phi U]^\dagger = XyZ^\dagger DZyX^\dagger = X[y^2 D ||z_i||^2_D]_d X^\dagger$. Mais alors que G-S est simple à faire, on a ici une SVD qui n'est pas simple — sauf à considérer un modèle quantique pour qui $\Phi = \exp(Ut)$ est unitaire ...

Bref, à part la factorisation LU, qui s'emboîte fondamentalement du fait que le produit matriciel conserve la trigonalisation ($LUU' = L[UU']$), le reste résiste résolument.

10 éléments du TEF de base

On rappelle la définition selon Cacuci de la différentielle de Gâteaux :

$$\mathcal{D}F(\eta; h) = \lim_{\epsilon \to 0} \frac{1}{\epsilon}(F(\eta + \epsilon h) - F(\eta))$$

pour ϵ réel et h arbitraire[1] ; et en pratique, Cacuci préconise d'analyser le terme supplémentaire $\Delta(h)$ dans la différence entre la $\mathcal{D}$ et la variation totale de F :

$$F(\eta + \epsilon h) - F(\eta) = \mathcal{D}F(\eta; h) + \Delta(h)$$

où on doit vérifier que : $\lim_{\epsilon \to 0} \Delta(h)/\epsilon = 0$. Pragmatiquement donc, je vois pas pourquoi on ne pourrais restreindre h à quelque sous-espace vectoriel présentant encore de l'intérêt — physiquement parlant.

On rappelle que dans la formulation de sensibilité d'une réponse (fonctionnelle) de Cacuci, la linéarisation d'un système en un point η n'est possible par différenciation de Gâteaux (notée $\mathcal{D}$) de $F(\eta)$ que ssi :

1. $F(\eta)$ satisfait une condition de Lipschitz faible, et
2. $F(\eta + \epsilon h_1 + \epsilon h_2) - F(\eta + \epsilon h_1) - F(\eta + \epsilon h_2) + F(\eta) = o(\epsilon)$

dans ce cas seulement, on note $\mathcal{D}F(\eta; h) = F'_e h$, oú F' est la dérivée de Gâteaux en η de F^2.

[1]je suppose que son existence impose $\forall h$?
[2]références en analyse fonctionnelle.

Lorsque les opérateurs d'évolution sont compliqués — comme dans le cas de l'utilisation d'une subroutine pour faire de la différenciation par différence finie — on doit assurer des critères permettant d'appliquer une technique de linéarisation ; il suffit peut-être de déterminer un ϵ intermédiaire entre la perte de précision par troncature et la perte de linéarité. Encore, l'espace vectoriel de h pourra être restreind mais physiquement pertinent.