Identification à l'aide de la transformée de Borel : Fit à 3 pôles et plus

(JYG-Mai 2003)

Le fit à 2 pôles s'avérant d'utilité très limitée, il est apparu de plus en plus nécessaire de développer un outil susceptible de fitter les résultats de simulations avec une fonction à trois pôles ou plus. Bien sûr, ce qui paraît indispensable pour le moment, c'est le fit à 3 pôles, mais on devine bien que l'on aura besoin d'aller plus loin bientôt. Et ce d'autant plus que, comme pour le fit à un pôle, il peut être nécessaire de commencer par un fit à un nombre pair de pôles.

Pour bien percevoir quelle est l'ampleur des difficultés, il faut réaliser que les objets élémentaires sont les pôles simples et les paires de pôles (à cause des pôles complexes conjugués). Et donc, ce qui a été fait jusqu'à présent, c'est de traiter les objets élémentaires : rien n'a été fait pour passer à des objets multiples. Le passage à des objets multiples va demander la mise au point d'une écriture générale d'une fonction à un nombre quelconque de pôles et la mise au point d'une stratégie de simulation, qui, à chaque étape, se ramènera à l'ajustement de paramètres variant continuement.

1 Ecriture générale d'une transformée Log-Borel à pôles réels

On considère une fonction concrète $g(t)$, solution générale d'un système d'E.D.O. linéaires :

$$g(t) = \sum_i (a_{i0} + a_{i1}t + a_{i2}t^2 + ... + a_{i(n_i-1)}t^{n_i-1})e^{-s_i t} \tag{1}$$

où s_i désigne le mode propre numéro i de multiplicité n_i et où les a_{ij} sont les amplitudes de ces modes. On suppose que $a_{i(n_i-1)}) \neq 0$,ce qui garantit que le mode i est de multiplicité n_i ; mais il faut noter que nous serons amenés plus loin à considérer tout l'espace vectoriel des fonctions de la forme (**??**), c'est-à-dire à considérer l'espace des fonctions de modes s_i avec des multiplicités **inférieures ou égales** aux n_i.

La transformé de Borel de $g(t)$ est alors :

$$\begin{aligned} \hat{g}(\tau) &= \sum_i (\frac{a_{i0}}{(1+s_i\tau)} + \frac{a_{i1}\tau}{(1+s_i\tau)^2} + \frac{2a_{i2}\tau^2}{(1+s_i\tau)^3} ... + \frac{(n_i-1)!a_{i(n_i-1)}\tau^{n_i-1}}{(1+s_i\tau)^{n_i}}) \\ &= \sum_i \sum_{j=0}^{n_i-1} \frac{j!a_{ij}\tau^j}{(1+s_i\tau)^{j+1}} \end{aligned} \tag{2}$$

Elle présente des pôles aux points $1/s_i$ avec les multiplicités n_i.

Développement en somme de termes relatifs à chaque pôle

En mettant au même dénominateur toutes les fractions relatives à un même pôle, on obtient une expression de la forme :

$$\hat{g}(\tau) = \sum_i \frac{P_{i,n_i-1}(\tau)}{(1+s_i\tau)^{n_i}} \tag{3}$$

où les P_{i,n_i-1} sont des polynomes de degré $n_i - 1$. La question est alors de savoir si la forme (**??**) est équivalente à (**??**).

En omettant tous les indices de pôle, le problème est donc de trouver les coefficients b_j tels que, pour un polynome $P_n(\tau) = \sum_{j=0}^{n} p_j \tau^j$ et pour un coefficient s donnés, l'égalité suivante soit satisfaite quel que soit τ :

$$\frac{P_n(\tau)}{(1+s\tau)^{n+1}} = \sum_{j=0}^{n} \frac{b_j \tau^j}{(1+s\tau)^{j+1}}$$

En multipliant les deux membres par $(1+s\tau)^{n+1}$ il vient :

$$\begin{aligned} P_n(\tau) &= \textstyle\sum_{j=0}^n b_j \tau^j (1+s\tau)^{n-j} \\ &= \textstyle\sum_{j=0}^n b_j \tau^j (\sum_{k=0}^{n-j} C_{n-j}^k s^k \tau^k)) \\ &= \textstyle\sum_{k=0}^n \tau^k [\sum_{j=0}^k b_j C_{n-j}^{k-j} s^{k-j}] \end{aligned} \tag{4}$$

Les équations donnant b_j sont donc :

$$\sum_{j=0}^{k} b_j C_{n-j}^{k-j} s^{k-j} = p_k \tag{5}$$

La matrice étant triangulaire, ce système se résoud simplement de proche en proche :

$$\begin{array}{l} b_0 = p_0 \\ b_1 + nsb_0 = p_1 \\ b_2 + (n-1)sb_1 + \frac{n(n-1)}{2}s^2 b_0 = p_2 \\ . \quad . \quad . \quad . \quad . \\ b_n + sb_{n-1} + s^2 b_{n-2} + ... + s^n b_0 = p_n \end{array} \tag{6}$$

L'expression (**??**) de $\hat{g}(\tau)$ est donc une forme générale, équivalente à (**??**). Il faut cependant noter que pour un polynome P_n de degré n, les b_i peuvent très bien être tous nuls à partir d'un certain rang (par exemple, si $P_n = (1+s\tau)^n$, alors $b_i = 0$ pour tout $i > 0$). La formule (**??**) représente donc les fonctions dont les pôles $1/s_i$ sont de multiplicités inférieures ou égales aux n_i (et non strictement égales aux n_i).

Développement en produit de termes relatifs à chaque pôle

Maintenant, on peut de nouveau regrouper les fractions de l'équation (??) en une seule fraction. Si on désigne par N la somme des multiplicités de tous les pôles, le dénominateur va être un polynome de degré N et le numérateur un polynome de degré $N-1$:

$$\hat{g}(\tau) = \frac{Q_{N-1}(\tau)}{\prod\limits_{i}(1+s_i\tau)^{n_i}} \tag{7}$$

où Q_{N-1} est un polynome de degré $N-1$.

Le passage de la forme (??) à la forme (??) se fait par décomposition de la fraction rationnelle qui constitue le membre de droite de (??) en éléments simples, puis par réduction au même dénominateur des fractions relatives au même pôle. Comme les dénominateurs sont tous normalisés de telle sorte que leur terme de degré 0 soit égal à 1, cette décomposition est unique. Il y a donc bijection entre les représentations (??) et (??) des transformées de Borel à pôles réels. Comme pour (??), on a ici une forme qui représente les fonctions ayant des pôles $1/s_i$ de multiplicités au plus n_i.

Log-Borel

La transformée Log-Borel s'obtient en posant $\tau = exp(\lambda)$ et $s_i = exp(-\rho_i)$ dans (??) :

$$\breve{g}(\lambda) = \frac{Q_{N-1}(e^\lambda)}{\prod\limits_{i}(1+e^{\lambda-\rho_i})^{n_i}} \tag{8}$$

2 Transformée Log-Borel à pôles quelconques

Dans le cas où des pôles complexes sont présents, apparaissant donc en paires de pôles complexes conjugués, on montre facilement que la forme (??) se généralise, avec un polynome Q_{N-1} à coefficients réels, en faisant apparaître explicitement les pôles complexes. Par exemple, si les pôles numéro 1 et 2 sont complexes conjugués et de multiplicité un, on aura :

$$\hat{g}(\tau) = \frac{Q_{N-1}(\tau)}{[1+(s_1+i\omega_1)\tau][1+(s_1-i\omega_1)\tau]\prod\limits_{i>2}(1+s_i\tau)^{n_i}}$$

Outre le fait que l'on va devoir traiter les pôles complexes par paire, il apparaît une difficulté liée au fait que l'on peut avoir besoin de passer continuement de deux pôles réels à un pôle double, puis à deux pôles complexes conjugués. Pour cela, il faut avoir une écriture commune à ces 3 cas, permettant de passer d'un cas à l'autre en variant un paramètre.

Ecriture unifiée de la partie du dénominateur de la fonction $\breve{g}$ concernant deux pôles voisins

On considère la partie D du dénominateur de $\breve{g}$ liée à deux pôles voisins qui peuvent être dans les configurations suivantes :

- deux pôles complexes conjugués : $s+i\omega = exp(-\rho+i\psi)$ et $s-i\omega = exp(-\rho-i\psi)$.
- un pôle double : $s = exp(-\rho)$.

- deux pôles réels espacés de δ : $s+\delta = exp(-\rho+\epsilon)$ et $s-\delta = exp(-\rho-\epsilon)$.

Un peu d'algèbre permet de montrer que D s'écrit suivant les cas :

- pôles complexes conjugués :

$$\begin{aligned} D &= (1+e^{\lambda-\rho-i\psi})(1+e^{\lambda-\rho+i\psi}) \\ &= 2e^{\lambda-\rho}[ch(\lambda-\rho)+cos\psi] \end{aligned} \tag{9}$$

- pôle double :

$$\begin{aligned} D &= (1+e^{\lambda-\rho})^2 \\ &= 2e^{\lambda-\rho}[ch(\lambda-\rho)+1] \end{aligned} \tag{10}$$

- deux pôles réels :

$$\begin{aligned} D &= (1+e^{\lambda-\rho-\epsilon})(1+e^{\lambda-\rho+\epsilon}) \\ &= 2e^{\lambda-\rho}[ch(\lambda-\rho)+ch\epsilon] \end{aligned} \tag{11}$$

Au voisinage d'un pôle double, on a donc une forme universelle dépendant continuement d'un paramètre réel ξ :

$$\left\{ \begin{array}{l} (1+e^{\lambda-\rho-\epsilon})(1+e^{\lambda-\rho+\epsilon}) = 2e^{\lambda-\rho}[ch(\lambda-\rho)+\xi] \\ \;\xi = 1 \quad : \epsilon = 0 \\ \;\xi < 1 \quad : \epsilon \text{ imaginaire} \\ \;\xi > 1 \quad : \epsilon \text{ réel} \end{array} \right. \tag{12}$$

Ecriture générale de la transformée Log-Borel à pôles quelconques

Une écriture générale est possible, grace aux deux propriétés :

- la formule (**??**) décrit une fonction dont les pôles sont de multiplicités inférieures ou égales aux n_i ; ceci implique en particulier qu'un pôle simple j peut figurer dans l'écriture (**??**) avec $n_j = 2$.
- on dispose d'un écriture générale du dénominateur de $\breve{g}$ pour les paires de pôles.

Pour profiter de ces propriétés, nous allons récrire la formule (**??**) (qui correspond au cas des pôles réels) en multipliant numérateur et dénominateur par le dénominateur :

$$\begin{aligned} \breve{g}(\lambda) &= \frac{Q_{N-1}(e^\lambda)}{\prod\limits_i (1+e^{\lambda-\rho_i})^{n_i}} \\ &= \frac{Q_{2N-1}(e^\lambda)}{\prod\limits_i [(1+e^{\lambda-\rho_i})^2]^{n_i}} \end{aligned} \tag{13}$$

ce qui peut s'écrire, en utilisant (**??**) :

$$\breve{g}(\lambda) = \frac{Q_{2N-1}(e^\lambda)}{\prod\limits_i (2e^{\lambda-\rho_i}[ch(\lambda-\rho_i)+1])^{n_i}} \tag{14}$$

Cette dernière formule permet de décrire tous les cas à pôles réels e^{ρ_i} de multiplicités inférieures ou égales à $2n_i$. Elle inclut en particulier les cas où les multiplicités sont simplement n_i (comme dans (??)).

On passe maintenant à une forme générale en substituant aux pôles doubles du dénominateur des paires de pôles (réels ou complexes conjugués) :

$$\breve{g}(\lambda) = \frac{Q_{2N-1}(e^{\lambda})}{\prod_i (2e^{\lambda-\rho_i}[ch(\lambda-\rho_i)+\xi_i])^{n_i}} \tag{15}$$

Cette formulation générale appelle plusieurs remarques :

- puisque les pôles y sont appariés, il peut y avoir plusieurs représentations d'une même fonction ; plus précisément, s'il y a p pôles réels, alors, il y a C_p^2 représentations de la fonction.
- les paires de pôles sont repérées par les couples (ρ_i, ξ_i).
- les pôles ayant $\xi_i = 1$ et $n_i = 1$ sont soit des pôles simples réels (et alors le polynome supérieur est un multiple de $1 + exp(\lambda - \rho_i)$), soit des pôles doubles (dans le cas contraire).
- dans la pratique, plutôt que de manipuler des exposants n_i, on va prendre uniquement des exposants unité, les mêmes ρ et ξ figurant eventuellement plusieurs fois dans les facteurs du dénominateur.

3 Ajustement par moindre χ^2

3.1 Formulation générale

On suppose que des simulations ont donné le graphe d'une fonction $\breve{f}(\lambda)$ sous la forme d'un échantillonage $\breve{f}_i = \breve{f}(\lambda_i)$. On se pose le problème de trouver une fonction $\breve{g}$, décrite sous la forme (??), qui fitte au mieux (à une constante près) le graphe de $\breve{f}$ sur un intervalle $[\Lambda_1, \Lambda_2]$. Ayant choisi le nombre N_0 de paires de pôles utilisées, on peut écrire $\breve{g}$:

$$\breve{g}(\lambda) = \frac{q_0 + q_1 e^{\lambda} + \cdots + q_{2N_0-1} e^{(2N_0-1)\lambda}}{\prod_{k=1}^{N_0} (2e^{\lambda-\rho_k}[ch(\lambda-\rho_k)+\xi_k])} \tag{16}$$

On cherche donc les paramètres $q_0,\ q_1, \ldots,\ q_{2N_0-1}$, puis $\rho_1,\ \rho_2, \ldots,\ \rho_{N_0}$ et, $\xi_1,\ \xi_2, \ldots,\ \xi_{N_0}$ et, enfin, C tels que le χ^2 suivant soit minimal :

$$\chi^2 = \sum_{\Lambda_1<\lambda_i<\Lambda_2} \left[\breve{f}_i - \frac{q_0 + q_1 e^{\lambda_i} + \cdots + q_{2N_0-1} e^{(2N_0-1)\lambda_i}}{\prod_{k=1}^{N_0} (2e^{\lambda_i-\rho_k}[ch(\lambda_i-\rho_k)+\xi_k])} - C \right]^2 \tag{17}$$

Le χ^2 peut aussi s'écrire :

$$\chi^2 = \sum_{\Lambda_1 < \lambda_i < \Lambda_2} \left[\check{f}_i - q_0 J_0(\lambda_i) - q_1 J_1(\lambda_i) + \cdots + q_{2N_0-1} J_{2N_0-1}(\lambda_i) - C\right]^2 \tag{18}$$

où l'on a posé :

$$\begin{aligned} J_j(\lambda) &= \frac{e^{j\lambda}}{\prod_{k=1}^{N_0} (2e^{\lambda-\rho_k}[ch(\lambda-\rho_k)+\xi_k])} \\ &= e^{j\lambda} J_0(\lambda) \end{aligned} \tag{19}$$

Pour procéder à la minimisation, on aura besoin des dérivées du vecteur des écarts par rapport aux $4N_0+1$ paramètres. En désignant par V_i les composantes de ce vecteur, on a :

$$\begin{aligned} \frac{\partial V_i}{\partial q_j} &= -J_j(\lambda_i) \\ \frac{\partial V_i}{\partial \rho_k} &= -q_0 \frac{\partial J_0(\lambda_i)}{\partial \rho_k} - \cdots - q_{2N_0-1} \frac{\partial J_{2N_0-1}(\lambda_i)}{\partial \rho_k} \\ &= -(q_0 + q_1 e^{\lambda_i} + \cdots + q_{2N_0-1} e^{(2N_0-1)\lambda_i}) \frac{\partial J_0(\lambda_i)}{\partial \rho_k} \\ \frac{\partial V_i}{\partial \xi_k} &= -q_0 \frac{\partial J_0(\lambda_i)}{\partial \xi_k} - \cdots - q_{2N_0-1} \frac{\partial J_{2N_0-1}(\lambda_i)}{\partial \xi_k} \\ &= -(q_0 + q_1 e^{\lambda_i} + \cdots + q_{2N_0-1} e^{(2N_0-1)\lambda_i}) \frac{\partial J_0(\lambda_i)}{\partial \xi_k} \\ \frac{\partial V_i}{\partial C} &= -1 \end{aligned} \tag{20}$$

Avec :

$$\begin{aligned} \frac{\partial J_0(\lambda)}{\partial \rho_k} &= J_0 \frac{e^{\lambda-\rho_k}+\xi_k}{ch(\lambda-\rho_k)+\xi_k} \quad \text{et} \quad \frac{\partial J_j(\lambda)}{\partial \rho_k} = e^{j\lambda} \frac{\partial J_0(\lambda)}{\partial \rho_k} \\ \frac{\partial J_0(\lambda)}{\partial \xi_k} &= J_0 \frac{1}{ch(\lambda-\rho_k)+\xi_k} \quad \text{et} \quad \frac{\partial J_j(\lambda)}{\partial \xi_k} = e^{j\lambda} \frac{\partial J_0(\lambda)}{\partial \xi_k} \end{aligned} \tag{21}$$

Avec ces dernières formules, on peut récrire les dérivées de façon plus compacte :

$$\begin{aligned} \frac{\partial V_i}{\partial \rho_k} &= -\check{g}(\lambda_i) \frac{e^{\lambda_i-\rho_k}+\xi_k}{ch(\lambda_i-\rho_k)+\xi_k} \\ \frac{\partial V_i}{\partial \xi_k} &= -\check{g}(\lambda_i) \frac{1}{ch(\lambda_i-\rho_k)+\xi_k} \end{aligned} \tag{22}$$

3.2 Mise en oeuvre