Université de Montréal
Faculté des arts et des sciences
Département de sciences économiques

CAHIER 9324

CAUSALITÉS À COURT ET À LONG TERME DANS LES MODÈLES VAR ET ARIMA MULTIVARIÉS

Jean-Marie Dufour[1] et Eric Renault[2]

[1] Département de sciences économiques et Centre de recherche et développement en économique (C.R.D.E.), Université de Montréal.

[2] GREMAQ et IDEI, Université des Sciences Sociales de Toulouse.

Août 1993

Ce travail a bénéficié de l'aide de la coopération franco-québécoise (projet « Analyse du changement structurel et prévision économique »), du Conseil de recherches en sciences humaines du Canada, du Conseil de recherches en sciences naturelles et en génie du Canada, du Fonds FCAR du Québec et de l'accueil des laboratoires C.R.D.E. (Montréal) et CREST-ENSAE (Paris). Nous désirons remercier John Galbraith, Eric Ghysels, Michel Mouchart et David Tessier pour leurs commentaires.

C.P. 6128, succursale A
Montréal (Québec)
H3C 3J7

Télécopieur (FAX): (514) 343-5831
Courrier électronique (E-Mail): econo@tornade.ERE.Umontreal.CA

Ce cahier a également été publié au Centre de recherche et développement en économique (C.R.D.E.) (publication no 1693).

Dépôt légal - 1993
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ISSN 0709-9231

ABSTRACT

Causality in the sense of Granger is typically defined in terms of predictibility of a vector of variables **one period ahead**. Recently, Lütkepohl (1990) proposed to define non-causality between two variables in terms of non-predictibility at any number of periods ahead. When more than two vectors are considered (i.e., when the information set contains auxiliary variables), these two notions are not equivalent. In this paper, we first generalize the notions of causality by considering causality at any given (arbitrary) horizon h. Then, we derive necessary and sufficient conditions for non-causality between vectors of variables (inside a larger vector) up to any given horizon h, where h can be infinite. The models considered include vector autoregressive processes, possibly of infinite order, and multivariate ARIMA processes. In particular, relatively simple *separation* and *rank* conditions for non-causality at any horizon h are derived.

Key words : ARIMA models; causality; forecast horizon; impulse response coefficients; multivariate time series model; rank condition; separation condition; vector autoregressive (VAR) model.

RÉSUMÉ

La causalité au sens de Granger est habituellement définie par la prévisibilité d'un vecteur de variables par un autre **une période à l'avance**. Récemment, Lütkepohl (1990) a proposé de définir la non-causalité entre deux variables (ou vecteurs) par la non-prévisibilité à tous les délais dans le futur. Lorsqu'on considère plus de deux vecteurs (i.e. lorsque l'ensemble d'information contient des variables auxiliaires), ces deux notions ne sont pas équivalentes. Dans ce texte, nous généralisons d'abord les notions antérieures de causalité en considérant la causalité à un horizon donné h arbitraire, fini ou infini. Ensuite, nous dérivons des conditions nécessaires et suffisantes de non-causalité entre deux vecteurs de variables (à l'intérieur d'un plus grand vecteur) jusqu'à un horizon donné h. Les modèles considérés incluent les autorégressions vectorielles, possiblement d'ordre infini, et les modèles ARIMA multivariés. En particulier, nous donnons des conditions de *séparabilité* et de *rang* pour la non-causalité jusqu'à un horizon h, lesquelles sont relativement simples à vérifier.

Mots clés : causalité; coefficients d'impulsion; condition de rang; condition de séparabilité; horizon de prévision; modèle ARIMA; modèle autorégressif; modèle de séries chronologiques multivarié.

1. INTRODUCTION

Le concept de causalité de Granger (1969) constitue une notion de base servant à étudier les liens dynamiques entre séries chronologiques. La littérature sur ce sujet est maintenant très vaste; voir, par exemple, Pierce et Haugh (1979), Newbold (1982), Geweke (1984), Gouriéroux et Monfort (1990, chapitre X) et Lütkepohl (1991). La définition originale de Granger (1969), qui est reprise par la plupart des auteurs sur ce sujet, réfère à la prévisibilité d'une variable $X(t)$, où t est un entier, à partir de son propre passé, de celui d'une autre variable $Y(t)$ et, possiblement, d'un vecteur de variables auxiliaires $Z(t)$, ***une période à l'avance*** : de façon plus précise, on dit que Y cause X au sens de Granger si l'observation de Y jusqu'au temps t $(Y(\tau) : \tau \leq t)$ peut aider à prévoir $X(t + 1)$, lorsque les observations correspondantes sur X et Z sont aussi disponibles $(X(\tau), Z(\tau) : \tau \leq t)$; une définition plus formelle sera fournie plus loin. Récemment, toutefois, Lütkepohl (1990) a fait remarquer que dans le cas d'un modèle multivarié où un vecteur de variables auxiliaires $Z(t)$ est utilisé en sus des variables d'intérêt $X(t)$ et $Y(t)$, il est possible que Y ne cause pas X en ce sens, mais puisse quand même aider à prévoir X plusieurs périodes à l'avance; sur cette question, voir aussi Sims (1980) ainsi que Renault et Szafarz (1991). Par exemple, les valeurs de $Y(\tau)$ jusqu'au temps t pourraient aider à prévoir $X(t + 2)$. Ceci est dû au fait que Y peut aider à prévoir Z une période à l'avance, qui à son tour a un effet sur X à une période subséquente. Il est clair que l'étude de tels effets indirects peut être d'un grand intérêt dans l'étude des relations entre séries chronologiques. En particulier, on peut ainsi faire la distinction entre des propriétés de « non-causalité à court terme » et de « non-causalité à long terme ».

Le but du présent article est d'étudier et caractériser de façon plus précise de tels effets. Dans la section 2, nous définissons d'abord des notions plus générales de causalité qui nous permettront d'aborder ces questions : causalité à un horizon donné h (où h est un entier positif) et causalité jusqu'à un horizon h quelconque (où $1 \leq h \leq \infty$). Ces définitions sont basées sur le concept de projection (causalité linéaire), ne requièrent pas la stationnarité des processus étudiés et, à l'horizon un $(h = 1)$, contiennent comme cas spécial la définition usuelle de la causalité au sens de Granger (1969). On peut ainsi étudier de façon précise des propriétés de « causalité à court terme » (h petit) et de « causalité à long terme » (h grand). Ensuite, nous présentons plusieurs résultats généraux sur la causalité jusqu'à un horizon quelconque. Nous donnons, en particulier, des caractérisations « composante par composante » des propriétés de causalité définies, ce qui permet de réduire l'étude des causalités entre vecteurs à celle de causalités entre variables scalaires, ainsi que des conditions suffisantes générales sous lesquelles la non-causalité à l'horizon 1 équivaut à la non-causalité à tous les horizons. Nous montrons que cette équivalence intervient dans deux cas importants : celui où les vecteurs X(t) et Y(t) contiennent toutes les variables considérées dans l'analyse, et celui où on peut « séparer » toutes les variables du système en deux sous-vecteurs qui ne se causent pas à l'horizon 1 (*condition de séparabilité*). Cette condition de séparabilité est équivalente à la définition de la « non-causalité » proposée par Hsiao (1982). Tous ces résultats sont démontrés pour des processus généraux dans L^2, sans hypothèse sur la stationnarité ou sur le caractère autorégressif (ou ARMA) de ces derniers.

On notera que la notion de non-causalité jusqu'à l'horizon infini étudiée ici n'est pas équivalente à celle considérée par Lütkepohl (1990). Cette dernière, en effet, n'est *pas* une généralisation de la définition usuelle de non-causalité au sens de Granger (1969), car la définition de Lütkepohl est basée sur l'absence d'effet des *innovations* d'une variable sur une autre variable (nullité des coefficients appropriés de la représentation moyenne mobile) plutôt que sur l'absence des valeurs passées d'une variable dans la prévision optimale d'une autre variable; sur ce point, voir Dufour et Tessier (1991).

Dans la section 3, nous étudions le cas où le processus considéré possède une représentation autorégressive possiblement d'ordre infini. Ces conditions incluent, bien sûr, comme cas spéciaux les processus autorégressifs d'ordre fini (VAR), stationnaires ou non stationnaires, une grande classe de processus stationnaires d'ordre deux, ainsi que les processus ARMA inversibles. Il n'est pas requis que la matrice de covariance des innovations soit constante (ce qui permet la présence d'hétéroscédasticité). Les résultats présentés ici généralisent de façon importante plusieurs résultats obtenus par Boudjellaba, Dufour et Roy (1992a, 1992b) pour l'horizon 1, ainsi que par Renault et Szafarz (1991) pour les processus multivariés autorégressifs d'ordre 1. Après avoir démontré une formule de récurrence qui permet de calculer de façon simple les prévisions à différents horizons, nous donnons une caractérisation des conditions de non-causalité à différents horizons en termes de « chaînes de causalité », formulation qui permet de mieux comprendre ces conditions. Dans la section 4, nous donnons des « conditions généralisées de séparabilité » qui sont suffisantes et intuitivement attrayantes. Dans le cas spécial où le système ne comporte qu'une seule variable auxiliaire, la condition de séparabilité est à la fois nécessaire et suffisante pour avoir non-causalité à tous les horizons. En d'autres termes, si le vecteur Z ne comporte qu'une seule composante, il n'y a que deux cas où Y ne cause pas X à tous les

horizons : celui où Y ne cause pas (X', Z)' à l'horizon 1 et celui où (Y', Z)' ne cause pas X à l'horizon 1 (on notera que X et Y demeurent des vecteurs). De plus, nous donnons des conditions de rang sur les matrices de la représentation autorégressive, conditions qui sont nécessaires pour avoir non-causalité à différents horizons. Ces dernières fournissent un moyen relativement simple de montrer qu'une hypothèse de non-causalité jusqu'à un horizon donné ne tient pas.

Dans la section 5, nous étudions le cas des processus ARIMA (non stationnaires) et montrons que la plupart des résultats démontrés pour les ARMA inversibles demeurent applicables pourvu que l'on fasse des hypothèses relativement faibles sur les conditions initiales.

2. NON-CAUSALITÉ LINÉAIRE À DIFFÉRENTS HORIZONS

Les concepts de non-causalité que nous considérons ici sont des extensions de la définition initiale de Granger (1969) dans un cadre ***linéaire*** conformément à Hosoya (1977) et Florens-Mouchart (1985). Plus précisément, on considère des définitions de la non-causalité qui s'expriment en termes de suites de conditions d'orthogonalité entre des sous-espaces vectoriels fermés d'un espace de Hilbert de variables aléatoires. Cet espace de Hilbert est un espace $L^2 = L^2(\Omega, \mathcal{A}, Q)$ de variables aléatoires réelles de carré intégrable, définies (à une égalité Q-presque sûre près) sur un espace de probabilité $(\Omega, \mathcal{A}, Q)$.

L'information (linéaire) « disponible à la date t » est dans ce contexte définie par un sous-espace de Hilbert $I(t)$ de L^2, où $t \in \mathbb{Z}$ et $\mathbb{Z}$ représente l'ensemble des entiers. Conformément à l'usage, on supposera que la suite $\{I(t) : t \in \mathbb{Z}, t > \omega\}$, où $\omega \in \mathbb{Z} \cup \{-\infty\}$, est croissante (au sens large), i.e.

$$(2.1) \qquad t < t' \Rightarrow I(t) \subseteq I(t') .$$

Autrement dit, la mémoire n'est pas bornée et on retient à la date t tout ce qui a été observé jusqu'à cette date. Notons, en particulier, que pour formaliser la notion d'environnement stationnaire, il peut être utile de considérer que l'observation a débuté en un passé aussi lointain que nécessaire; l'ensemble des dates d'observation est alors l'ensemble des entiers ($\omega = -\infty$).

En particulier, il est usuel de dire, en termes heuristiques, que l'on dispose à la date t de l'observation jusqu'à la date t d'un processus d'intérêt $\{X(\tau) : \tau \in \mathbb{Z}, \tau > \omega\}$. Plus précisément, $X(\tau)$ désigne un vecteur de m variables réelles $x_i(\tau)$, $i = 1, 2, \ldots, m$, et on suppose que les composantes $x_i(\tau)$ de $X(\tau)$ appartiennent à $I(t)$ pour $\omega < \tau \leq t$.

Définition 1 : La suite croissante de sous-espaces de Hilbert $\{I(t) : t \in \mathbb{Z}, t > \omega\}$ de L^2 est dite ***conforme au processus*** $\{X(t) : t \in \mathbb{Z}, t > \omega\}$ si pour tout entier $t > \omega$,

(2.2) $\quad X(\omega, t] \subseteq I(t)$

où $X(\omega, t]$ désigne l'espace de Hilbert engendré par les composantes $x_i(\tau)$, $i = 1, 2, \ldots, m$, de $X(\tau)$, $\omega < \tau \leq t$.

La valeur de ω n'est pas précisée à ce stade : elle peut être en particulier $-\infty$ ou 0 selon que l'on s'intéresse à la dynamique d'un processus stationnaire sur les entiers ($t \in \mathbb{Z}$) ou à une dynamique de $X(t)$, $t \geq 1$, conditionnelle à des valeurs initiales antérieures à la date 1. L'ensemble $I(\omega) = \bigcap_{t > \omega} I(t)$ représente l'information disponible à toute date $t > \omega$, telles des constantes, des variables déterministes ou des conditions initiales (relatives à $X(t)$ ou à d'autres variables).

Quoiqu'il en soit, la connaissance de $I(t)$ (et en particulier de $X(\tau)$, $\omega < \tau \leq t$) ne permet pas, en général, de prévoir parfaitement une valeur future $X(t + h)$, où $h \in \mathbb{N}$ désigne l'*horizon de la prévision*; $\mathbb{N} = \{1, 2, \ldots\}$ représente l'ensemble des entiers positifs tandis que $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ est l'ensemble des entiers non négatifs. La ***meilleure prévision linéaire*** de $X(t + h)$, étant donné l'information $I(t)$ sur l'« univers » environnant est notée $P[X(t + h) \mid I(t)]$ et constitue une variable aléatoire de dimension m dont les composantes sont les projections orthogonales $P[x_i(t + h) \mid I(t)]$ des composantes $x_i(t + h)$, $i = 1, 2, \ldots, m$, de $X(t)$ sur le sous-espace de Hilbert $I(t)$ (projections orthogonales au sens du produit scalaire usuel de L^2). Il s'agit en fait de la ***régression affine de*** $X(t + h)$ ***sur*** $I(t)$ (qui coïncide avec l'espérance conditionnelle dans le cas gaussien) si l'on suppose que l'information minimale $I(\omega)$ contient en particulier une variable constante non nulle.

L'idée des concepts de causalité (linéaire) au sens de Granger d'un processus Y ***vers un processus*** X consiste à étudier l'amélioration de la prévision de $X(t + h)$ qui

peut être apportée par la connaissance à la date t, outre de $I(t)$, de l'ensemble des valeurs présentes et passées $Y(\tau)$, $\omega < \tau \leq t$, du processus Y. Y étant un vecteur de dimension p, $Y(t) = (y_1(t), \ldots, y_p(t))'$, on notera $I(t) + Y(\omega, t]$ le sous-espace de Hilbert dans L^2 engendré par $I(t)$ et les composantes $y_j(\tau)$, $j = 1, \ldots, p$, $\omega < \tau \leq t$. On introduit alors des concepts de causalité qui sont, évidemment, ***relatifs à l'« univers »*** $I(t)$, même si la notation ne rappelle pas que la propriété dépend crucialement de l'univers considéré.

Définition 2 : Étant donné une suite « d'univers » $I = \{I(t) : t \in \mathbb{Z}, t > \omega\}$ conforme au processus $X = \{X(t) : t \in \mathbb{Z}, t > \omega\}$ et $h \in \mathbb{N}$, on dit que :

(i) Y ne cause pas X *à l'horizon h* (noté $Y \underset{h}{\not\rightarrow} X$) relativement à l'univers I, si

(2.3) $\quad P[X(t + h) \mid I(t)] = P[X(t + h) \mid I(t) + Y(\omega, t]]\ , \forall t > \omega\ ;$

(ii) Y ne cause pas X *jusqu'à l'horizon h* (noté $Y \underset{(h)}{\not\rightarrow} X$) relativement à l'univers I si $Y \underset{k}{\not\rightarrow} X$ pour $k = 1, 2, \ldots, h$;

(iii) Y ne cause pas X *à tout horizon* (noté $Y \underset{\infty}{\not\rightarrow} X$) relativement à I si $Y \underset{k}{\not\rightarrow} X$ relativement à I pour tout $k \in \mathbb{N}$.

Par rapport à la définition usuelle de la non-causalité au sens de Granger (1969), la définition 2 appelle trois commentaires :

- d'une part, le rôle de l'horizon h de la prévision (où $h \in \mathbb{N}$) dans l'analyse de causalité est souligné. Comme l'ont remarqué récemment Lütkepohl (1990) et Renault-Szafarz (1991), la propriété de non-causalité, habituellement considérée à l'horizon 1 (égal à l'intervalle entre les observations) n'est en général ni une condition nécessaire, ni une condition suffisante pour que l'information $Y(\omega, t]$ soit inutile dans la prévision de X à un horizon $h \geq 2$;

- d'autre part, l'éventuelle non-stationnarité des processus considérés implique que la prévision à l'horizon h soit étudiée *à chaque date t*, la propriété (2.3) pour une date t donnée n'impliquant pas en général la même propriété pour une autre date;

– enfin, l'univers au sein duquel sont analysées les propriétés de causalité de Y à la date t contient, outre l'observation du processus d'intérêt X jusqu'à la date t, une information $I(\omega) = \bigcap_{t>\omega} I(t)$ disponible à toute date (qui peut regrouper en particulier un certain nombre de conditions initiales relatives à plusieurs variables) et un *supplément d'information* (différence entre I(t) et $I(\omega)$) qui représente l'information accumulée par l'observation entre ω et la date t, non seulement du processus d'intérêt X, mais aussi éventuellement de tierces variables composantes d'un processus $\{Z(t) : t \in \mathbb{Z}, t > \omega\}$. Il est clair aussi que le choix de la date initiale ω peut affecter les propriétés de causalité des processus considérés.

C'est précisément la présence de ces tierces variables qui implique, comme on le constatera dans les sections suivantes, qu'une causalité de Y vers X, absente jusqu'à l'horizon h, peut apparaître à l'horizon h + 1. Sans détailler davantage les tierces variables contenues dans l'univers I, il est déjà possible d'énoncer quelques résultats de base (propositions 1 à 3 ci-dessous) dont on peut retrouver des versions équivalentes dans divers contextes particuliers examinés dans la littérature; voir, par exemple, Florens et Mouchart (1985) et Boudjellaba, Dufour et Roy (1992a).

Proposition 1 : Étant donné un univers I conforme au processus X de composantes x_i, $i = 1, 2, \ldots, m$, i.e. $X(t) = (x_1(t), \ldots, x_m(t))'$, les deux propriétés suivantes sont équivalentes pour tout $h \in \mathbb{N}$:

(i) $Y \underset{h}{\not\rightarrow} X$;

(ii) $Y \underset{h}{\not\rightarrow} x_i$ pour $i = 1, 2, \ldots, m$.

Cette proposition est évidente, compte tenu de la définition de la prévision $P[X(t+h) \mid I(t)]$ comme vecteur des prévisions $P[x_i(t+h)' \mid I(t)]$, $i = 1, 2, \ldots, m$. On prendra garde, cependant, à ce que pour chaque $i = 1, 2, \ldots, m$, la propriété de non-causalité $Y \underset{h}{\not\rightarrow} x_i$ soit considérée relativement à un univers I *conforme à* X *et non pas seulement à* x_i. La proposition 1 pourrait, en effet, être fausse [comme l'ont prouvé Florens et Mouchart (1985, Property 3.5, p. 164)] si l'on définissait la propriété $Y \underset{h}{\not\rightarrow} x_i$ à partir d'un univers dépendant de la composante x_i considérée. Les propositions 2 et 3 ci-dessous (relatives, quant à elles, aux composantes de Y) complètent la question de la caractérisation « *composante par composante* » de la causalité de Y vers X.

Proposition 2 : Étant donné un univers I conforme au processus X et un processus Y de composantes y_j, $j = 1, 2, \ldots, p$, i.e. $Y(t) = (y_1(t), \ldots, y_p(t))'$, soit $I_{)j(}(t)$ l'espace de Hilbert engendré par I(t) et les variables $y_k(\tau)$, $\omega < \tau \le t$, $k = 1, 2, \ldots, p$, $k \ne j$. Pour tout $h \in \mathbb{N}$,

(i) si $Y \underset{h}{\not\rightarrow} X$ relativement à l'univers I, alors pour $j = 1, 2, \ldots, p$,

(2.4) $P[X(t+h) \mid I_{)j(}(t)] = P[X(t+h) \mid I(t) + Y(\omega, t]]$, $\forall t > \omega$;

(ii) la réciproque de l'implication (i) est fausse en général.

Les démonstrations des propositions et des théorèmes sont fournies en annexe. L'identité (2.4) signifie que $y_j \not\rightarrow X$ relativement à $I_{)j(}$, pour $j = 1, 2, \ldots, p$: *à partir de l'information complète* $I(t) + Y(\omega, t]$, on peut, sans dommage pour la qualité de la prévision de X(t + h), se passer de l'information fournie par la composante y_j (ceci pour chaque $j = 1, 2, \ldots, p$). (2.4) n'implique (2.3) que si les composantes de Y ont certaines propriétés d'indépendance linéaire (comme on le vérifiera pour des modèles particuliers dans les sections suivantes). En revanche, si *à partir de l'information minimum* I(t) on peut sans dommage se passer de l'information fournie par chaque composante y_j (pour $j = 1, 2, \ldots, p$), alors on peut affirmer que Y ne cause pas X. Ce dernier fait est établi dans la proposition suivante.

Proposition 3 (Caractérisation composante par composante de la causalité entre vecteurs) : Étant donné un univers I conforme au processus X et un processus Y de composantes y_j, $j = 1, 2, \ldots, p$, les trois propriétés suivantes sont équivalentes pour tout $h \in \mathbb{N}$:

(i) $Y \underset{h}{\not\rightarrow} X$ relativement à I ;

(ii) $y_j \underset{h}{\not\rightarrow} X$ relativement à I , pour $j = 1, 2, \ldots, p$;

(iii) $y_j \underset{h}{\not\rightarrow} x_i$ relativement à I, pour $i = 1, 2, \ldots, m$ et $j = 1, 2, \ldots, p$.

Cette possibilité de caractériser la non-causalité de Y vers X *composante par composante à univers fixé* (conjonction des propositions 1 et 3) est une propriété agréable de la non-causalité *linéaire*, l'argument d'algèbre linéaire (en termes d'orthogonalité) employé ci-dessus ne pouvant être étendu à des concepts d'indépendance conditionnelle. Le concept probabiliste sous-jacent aux relations d'orthogonalité considérées ci-dessus est en fait celui de *non-corrélation conditionnelle* dès que l'on considère que les variables d'intérêt sont centrées ou que l'espace de Hilbert minimal $I(\omega)$ contient en particulier les variables constantes. Enfin, comme annoncé ci-dessus, on démontre, en annexe, la proposition 4 suivante qui prouve que le problème de l'horizon de la causalité ne se pose réellement que si l'univers contient d'autres processus que les processus X et Y mis en jeu dans la relation de causalité.

Proposition 4 (Condition suffisante d'équivalence entre les non-causalités à différents horizons) : Étant donné un sous-espace de Hilbert E de L^2 et un processus $X = \{X(t) : t \in \mathbb{Z},\ t > \omega\}$, où $X(t) = (x_1(t), \dots, x_m(t))'$ avec $x_i(t) \in L^2$, considérons l'univers conforme à X défini par $I(t) = E + X(\omega, t]$ pour $t > \omega$. Alors, pour tout processus $Y = \{Y(t) : t \in \mathbb{Z},\ t > \omega\}$, où $Y(t) = (y_1(t), \dots, y_p(t))'$ avec $y_j(t) \in L^2$, les propriétés suivantes sont équivalentes :

(i) $Y \underset{1}{\not\rightarrow} X$ relativement à I ;

(ii) $Y \underset{(h)}{\not\rightarrow} X$ relativement à I, $\forall h \in \mathbb{N}$;

(iii) $Y \underset{\infty}{\not\rightarrow} X$ relativement à I.

La proposition 4 donne donc un cas où la notion usuelle de non-causalité, au sens de Granger ($Y \underset{1}{\not\rightarrow} X$), implique effectivement que la connaissance du passé de Y n'est *jamais* utile pour prévoir X, i.e. $Y \underset{\infty}{\not\rightarrow} X$. C'est le cas où la seule information susceptible de s'ajouter à I(t), lorsque t grandit, est celle contenue dans X(t) et Y(t).

Considérons maintenant le cas d'un univers plus riche que celui de la proposition 4 parce qu'il contient aussi l'observation, à chaque date t, de variables réelles $Z_k(t)$ dans L^2, $k = 1, 2, \dots, r$, composantes d'un processus Z(t). On notera dans ce cas l'information de référence à la date t :

(2.5) $$I_{XZ}(t) = E + X(\omega, t] + Z(\omega, t]\,.$$

Alors la proposition 4 n'est plus directement applicable pour caractériser les cas où Y *ne cause pas* X *à tout horizon relativement à l'univers* I_{XZ}. Cependant, un cas particulier important où l'on peut montrer que Y ne cause jamais X est celui où une condition de *séparabilité* se trouve satisfaite.

Supposons, en effet, qu'il existe deux processus $\{Z_1(t) : t \in \mathbb{Z},\ t > \omega\}$ et $\{Z_2(t) : t \in \mathbb{Z},\ t > \omega\}$ de dimensions respectives r_1 et r_2 tels que, à chaque date t, l'univers $I_{XZ}(t)$ soit engendré par la réunion de

(2.6) $$I_{XZ_1}(t) = E + X(\omega, t] + Z_1(\omega, t]$$

et de $Z_2(\omega, t]$:

(2.7) $$I_{XZ}(t) = I_{XZ_1}(t) + Z_2(\omega, t] = I_{XZ_1Z_2}(t)\,.$$

Il est clair que ceci est en particulier vérifié dans le cas où le vecteur aléatoire $(Z_1(t)', Z_2(t)')'$ est une fonction linéaire bijective de Z(t). Par abus d'écriture, on considérera que r_1 ou r_2 peuvent être nuls, cas extrêmes qui correspondent respectivement à $Z = Z_2$ ou $Z = Z_1$.

Supposons, en outre, que $\begin{bmatrix} Y \\ Z_2 \end{bmatrix} \underset{1}{\not\rightarrow} \begin{bmatrix} X \\ Z_1 \end{bmatrix}$ relativement à l'univers I_{XZ_1}. Alors, d'après la proposition 4, on peut affirmer que

$$\begin{bmatrix} Y \\ Z_2 \end{bmatrix} \underset{\infty}{\not\rightarrow} \begin{bmatrix} X \\ Z_1 \end{bmatrix} \text{ relativement à } I_{XZ_1}\,.$$

En particulier, d'après la proposition 1 ,

$$\begin{bmatrix} Y \\ Z_2 \end{bmatrix} \underset{\infty}{\not\rightarrow} X \text{ relativement à } I_{XZ_1}\,.$$

En d'autres termes, pour tout entier positif h,

$$P[X(t+h) \mid I_{XZ_1}(t) + Z_2(\omega, t] + Y(\omega, t]] = P[X(t+h) \mid I_{XZ_1}(t)]\,,\ \forall t > \omega\,,$$

ou encore

$$P[X(t+h) \mid I_{XZ}(t) + Y(\omega, t]] = P[X(t+h) \mid I_{XZ_1}(t)]\,,\ \forall t > \omega\,.$$

Ceci prouve que chaque composante de $P[X(t+h) \mid I_{XZ}(t) + Y(\omega, t]]$ est un élément de $I_{XZ_1}(t) \subseteq I_{XZ}(t)$ pour $t > \omega$, et donc *a fortiori* que

$$P[X(t+h) \mid I_{XZ}(t) + Y(\omega, t]] = P[X(t+h) \mid I_{XZ}(t)], \forall t > \omega,$$

c'est-à-dire que $Y \underset{\infty}{\not\to} X$ relativement à I_{XZ}. Nous avons donc démontré la proposition suivante.

Proposition 5 (Condition de séparabilité pour la non-causalité à tous les horizons) : Supposons que E un sous-espace de Hilbert de L^2, $X(t) = (x_1(t), \dots, x_m(t))'$, $Y(t) = (y_1(t), \dots, y_p(t))'$ et $Z(t) = (z_1(t), \dots, z_r(t))'$, où $t \in \mathbb{Z}$ et $t > \omega$, sont trois processus dont les composantes x_i, y_j, z_k sont des éléments de L^2, tandis que Z_1 et Z_2 sont deux processus tels que $I_{XZ}(t)$ coïncide avec $I_{XZ_1Z_2}(t)$ pour tout entier $t > \omega$. Alors, la *condition de séparabilité*

$$\begin{bmatrix} Y \\ Z_2 \end{bmatrix} \underset{1}{\not\to} \begin{bmatrix} X \\ Z_1 \end{bmatrix} \text{ relativement à } I_{XZ_1}$$

est une condition suffisante pour que $Y \underset{\infty}{\not\to} X$ relativement à I_{XZ}.

La proposition 5 signifie intuitivement que, lorsqu'il y a séparabilité, non seulement $Y \underset{1}{\not\to} X$, mais de plus, Y *ne cause jamais* X. Ceci provient du fait qu'aucune « chaîne de causalité entre Y et X » [voir Renault et Szafarz (1991)] ne peut se constituer en transitant par Z (causalité indirecte), car les transformations linéaires de Z susceptibles d'être causées par Y (les composantes de Z_2) ne causent pas X.

L'intérêt des modèles de processus « linéaires » ARMA et ARIMA que nous allons considérer dans les sections suivantes est de fournir des formulations explicites des chaînes de causalité (par produit de matrices) permettant de caractériser avec plus de précision (en dehors de la condition suffisante fournie par la proposition 5 ci-dessus) les cas où Y ne cause jamais X et, plus généralement, où Y ne cause pas X jusqu'à un horizon donné.

3. NON-CAUSALITÉS DANS UN PROCESSUS LINÉAIRE INVERSIBLE

On supposera, dans cette section, que l'information I(t) est définie par l'observation jusqu'à la date t de deux processus : le processus X(t) d'intérêt (de dimension m) et un autre processus Z(t) (de dimension r), où $t \in \mathbb{Z}$. *L'information de référence* est :

$$(3.1) \qquad I(t) = I_{XZ}(t) = E + X(-\infty, t] + Z(-\infty, t],$$

où $t > \omega$, $\omega \in \mathbb{Z} \cup \{-\infty\}$ et E inclut une variable constante non nulle (et possiblement d'autres variables connues à toute date $t > \omega$). Cette spécification correspond à notre choix d'étudier ici la dynamique conjointe de trois processus : processus X susceptible d'être causé, processus Y susceptible de causer et processus Z complétant la description de « l'univers ». On sera parfois amené à considérer aussi les propriétés de causalité par rapport aux univers I_{XY} ou I_{YZ}, mais *l'univers de référence* sera I_{XZ} dans la mesure où l'on cherche finalement à caractériser les non-causalités de Y *vers* X *sachant que* Z *est dans l'univers*. De plus, le fait que I(t) inclue une constante signifie que nous considérons des *projections affines*. On notera enfin que, même si les processus X(t), Y(t) et Z(t) sont supposés définis pour tout $t \in \mathbb{Z}$, la propriété de non-causalité est toujours définie par rapport à une date initiale ω [voir (2.3)]. Pour que l'identité (2.3) tienne pour tout $t \in \mathbb{Z}$, il suffit de prendre $\omega = -\infty$. Ce formalisme nous permettra d'étudier dans la même foulée tant des processus stationnaires que non-stationnaires (avec des conditions initiales).

Pour faciliter le calcul des projections sur les sous-espaces I(t) et $I(t) + X(-\infty, t]$, nous allons faire l'hypothèse supplémentaire suivante. Pour $t > \omega$, où $\omega \in \mathbb{Z} \cup \{-\infty\}$, le processus $\{W(t) : t \in \mathbb{Z}\}$, où $W(t) = (X(t)', Y(t)', Z(t)')'$, possède une représentation autorégressive infinie, i.e.

$$(3.2) \qquad W(t) = \sum_{k=1}^{\infty} \pi_k W(t-k) + a(t), \forall t > \omega,$$

où $\{a(t) : t \in \mathbb{Z}\}$ est une suite de vecteurs aléatoires dans L^2 de moyenne nulle, non corrélés entre eux et tels que a(t) est orthogonal à l'espace de Hilbert $I(t-1) + Y(-\infty, t-1]$; de plus, nous supposerons que la série $\sum_{k=1}^{\infty} \pi_k W(t-k)$ converge en moyenne quadratique pour tout $t > \omega$. Il n'est pas requis que la matrice de covariance de a(t) soit constante. Par conséquent, a(t) peut ne pas être un bruit blanc. Lorsque la condition (3.2) sera satisfaite, nous dirons que le processus $\{W(t) : t \in \mathbb{Z}, t > \omega\}$ vérifie (3.2).

Les résultats sur la causalité obtenus sur la base du modèle (3.2) seront aussi applicables au modèle plus général

$$X(t) = \mu(t) + \sum_{k=1}^{\infty} \pi_k \, W(t-k) + a(t) \, , \, t > \omega \, , \tag{3.3}$$

où $\mu(t) = [\mu_1(t), \ldots, \mu_m(t)]'$, et $\mu_i(t) \in E$ pour tout $t > \omega$ et $i = 1, \ldots, m$ (e.g., $\mu(t)$ est une constante ou une « fonction de tendance » déterministe), ou encore au modèle

$$X(t) - \bar{\mu}(t) = \sum_{k=1}^{\infty} \pi_k \, [W(t-k) - \bar{\mu}(t-k)] \, , \, t > \omega \, , \tag{3.4}$$

où $\bar{\mu}_i(t) \in E$, $i = 1, \ldots, m$, pour tout $t \in \mathbb{Z}$ et la série $\sum_{k=1}^{\infty} \pi_k \, \bar{\mu}(t-k)$ converge pour tout $t > \omega$. Pour ce faire, il suffira que l'ensemble E inclue les variables appropriées, telle une constante non nulle lorsque $\mu(t)$ ou $\bar{\mu}(t)$ sont des constantes. Le fait de considérer le modèle (3.2) plutôt que (3.3) ou (3.4) permettra de simplifier notre exposé sans perte de généralité. En outre, dans le cas où (3.2) s'applique directement, on peut aussi poser que l'espace E est vide, i.e. $I(t) = X(-\infty, t] + Z(-\infty, t]$, sans affecter la validité des résultats énoncés dans cette section et la suivante.

Un cas important où les modèles (3.2) ou (3.4) s'appliquent est celui d'un processus stationnaire possédant une représentation autorégressive (possiblement d'ordre infini). Dans ce cas, on supposera aussi que les matrices carrées π_k sont absolument sommables, i.e.

$$\sum_{k=1}^{\infty} ||\pi_k|| < \infty \tag{3.5}$$

où $||\pi_k||^2 = \mathrm{tr}(\pi_k \pi_k')$. Un autre cas important où (3.2)–(3.4) s'appliquent est celui d'un processus ARIMA inversible. La forme autorégressive (3.2) fournit naturellement le calcul des prévisions à n'importe quel horizon h compte tenu de l'information $I(t) + Y(-\infty, t]$ disponible à la date t. On vérifie alors facilement la proposition suivante.

Proposition 6 (Formules de récurrence pour les coefficients des projections) : Si le processus $\{W(t) : t \in \mathbb{Z}, t > \omega\}$ vérifie (3.2), alors

$$P[W(t+h) \mid I(t) + Y(-\infty, t]] = \sum_{j=1}^{\infty} \pi_j^{(h)} \, W(t+1-j) \, , \, \forall t > \omega \, , \, \forall h \in \mathbb{N} \, , \tag{3.6}$$

où, pour chaque j, la suite de matrices $\pi_j^{(h)}$, $h \in \mathbb{N}$, est définie par récurrence sur h :

$$\pi_j^{(1)} = \pi_j \, , \quad \pi_j^{(h+1)} = \pi_{j+h} + \sum_{\ell=1}^{h} \pi_{h-\ell+1} \, \pi_j^{(\ell)} \, , \, h = 1, 2, \ldots \tag{3.7}$$

De plus, toute suite de matrices $\pi_j^{(h)}$, où $j \in \mathbb{N}$ et $h \in \mathbb{N}$, qui satisfait (3.7) satisfait aussi les relations de récurrence :

$$\pi_j^{(h+1)} = \pi_{j+1}^{(h)} + \pi_1^{(h)} \, \pi_j \, , \, h = 1, 2, \ldots \tag{3.8}$$

La proposition 6 permet de caractériser les propriétés de non-causalité $Y \underset{h}{\not\rightarrow} X$ relativement à l'univers I à partir de partitions naturelles des matrices $\pi_j^{(h)}$ définies par (3.7) selon X, Y et Z :

$$\pi_j^{(h)} = \begin{bmatrix} \pi_{XXj}^{(h)} & \pi_{XYj}^{(h)} & \pi_{XZj}^{(h)} \\ \pi_{YXj}^{(h)} & \pi_{YYj}^{(h)} & \pi_{YZj}^{(h)} \\ \pi_{ZXj}^{(h)} & \pi_{ZYj}^{(h)} & \pi_{ZZj}^{(h)} \end{bmatrix} . \tag{3.9}$$

On montre le résultat suivant qui est une généralisation de la proposition 1 de Boudjellaba, Dufour et Roy (1990).

Proposition 7 : Si le processus $\{W(t) : t \in \mathbb{Z}, t > \omega\}$ vérifie (3.2), la condition

$$\pi_{XYj}^{(h)} = 0 \, , \, j = 1, 2, \ldots \tag{3.10}$$

est suffisante pour que $Y \underset{h}{\not\rightarrow} X$ relativement à I_{XZ}. Si, de plus, la matrice de covariance $E[a(t)\,a(t)']$ est non singulière pour tout $t > \omega$, la condition (3.10) est nécessaire et suffisante pour que $Y \underset{h}{\not\rightarrow} X$ relativement à I_{XZ}.

La conjonction des propositions 6 et 7 permet de mieux comprendre pourquoi il est possible, en présence de tierces variables Z, que Y cause X à l'horizon h + 1, alors que $Y \underset{h}{\not\longrightarrow} X$. En effet, la relation de récurrence (3.8) implique que

$$\pi_{XYj}^{(h+1)} = \pi_{XY,j+1}^{(h)} + \pi_{XX1}^{(h)} \pi_{XYj} + \pi_{XY1}^{(h)} \pi_{YYj} + \pi_{XZ1}^{(h)} \pi_{ZYj} .$$

En appliquant la proposition 7, on trouve alors le résultat suivant.

Proposition 8 : Sous le modèle (3.2), si $Y \underset{(h)}{\not\longrightarrow} X$ relativement à I_{XZ} et si les matrices de covariance $E[a(t)\, a(t)']$ sont non singulières pour tout $t > \omega$, alors

$$\pi_{XYj}^{(h+1)} = \pi_{XZ1}^{(h)} \pi_{ZYj} , \; j = 1, 2, \ldots .$$

Autrement dit, c'est parce que les tierces variables Z sont susceptibles de causer X à l'horizon h, i.e. $\pi_{XZ1}^{(h)} \neq 0$, et qu'elles sont elles-mêmes causées par Y à l'horizon 1 ($\pi_{ZYj} \neq 0$) que peut apparaître une causalité de Y vers X à l'horizon h + 1, alors qu'elle n'est pas apparue jusqu'à l'horizon h. Ainsi, la présence de Z peut introduire une causalité indirecte de Y vers X passant par Z. Ceci permet évidemment d'écrire des *conditions suffisantes* de non-causalité de Y vers X, généralisant la proposition 5. En effet si, par exemple, $Z = (Z_1', Z_2')'$, on voit que

$$Y \underset{(h)}{\not\longrightarrow} X \Rightarrow \pi_{XYj}^{(h+1)} = \pi_{XZ_1 1}^{(h)} \pi_{Z_1 Yj} + \pi_{XZ_2 1}^{(h)} \pi_{Z_2 Yj} ,$$

d'où des conditions suffisantes du type :

$$Y \underset{1}{\not\longrightarrow} \begin{bmatrix} X \\ Z_1 \end{bmatrix} \text{ et } Z_2 \underset{(h)}{\not\longrightarrow} X \Rightarrow Y \underset{(h+1)}{\not\longrightarrow} X ,$$

que l'on laisse le soin au lecteur de vérifier (avec des définitions convenables des univers associés aux différentes relations de non-causalité). Pour réaliser une étude plus systématique, notre résultat essentiel est le théorème 1 suivant et ses corollaires.

***Théorème 1** (Caractérisation par chaînes de la non-causalité jusqu'à un horizon h)* : Soit $\{W(t) : t \in \mathbb{Z}, t > \omega\}$ un processus qui vérifie (3.2) et supposons que les matrices de covariance $E[a(t)\, a(t)']$ sont non singulières pour tout $t > \omega$. Alors, pour tout entier h supérieur à 1, les trois propriétés suivantes sont équivalentes :

(i) $Y \underset{(h)}{\not\longrightarrow} X$ relativement à I_{XZ} ;

(ii) $\pi_{XYj} = 0$, $\forall j \geq 1$, et $\pi_{XZ1}^{(k)} \pi_{ZYj} = 0$, $k = 1, 2, \ldots, h-1$, $\forall j \geq 1$;

(iii) $\pi_{XYj} = 0$, $\forall j \geq 1$, et $R_{XZ}^{(k)} \pi_{ZYj} = 0$, $k = 1, 2, \ldots, h-1$, $\forall j \geq 1$

où

$$R_{XZ}^{(k)} = \sum_{\ell=1}^{k} \pi_{XZ\ell} \left\{ \sum_{J(k-\ell)} \left[\prod_{i=1}^{k-\ell} \pi_{ZZi}^{n_i} \right] \right\} ,$$

$$J(\ell) = \left\{ (n_1, n_2, \ldots, n_\ell) : \sum_{i=1}^{\ell} i n_i = \ell \text{ et } n_i \in \mathbb{N}_0 , \; i = 1, \ldots, \ell \right\} ,$$

$$\prod_{i=1}^{k-\ell} \pi_{ZZi}^{n_i} = \pi_{ZZ1}^{n_1} \pi_{ZZ2}^{n_2} \cdots \pi_{ZZ,k-\ell}^{n_{k-\ell}}$$

avec, par convention, $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ et $\sum_{J(0)} \left[\Pi \, \pi_{ZZi}^{n_i} \right] = I_r$.

L'intérêt des critères (ii) et (iii) de non-causalité $Y \underset{(h)}{\not\longrightarrow} X$ est à apprécier par comparaison avec celui fourni par la proposition 7 : $\pi_{XYj}^{(k)} = 0$, $k = 1, 2, \ldots, h$, $\forall j \geq 1$. Le critère (ii) présente l'inconvénient de s'appuyer sur les matrices $\pi_{XZ1}^{(k)}$, $k = 1, 2, \ldots, h-1$, dont la formule n'est pas plus explicite que celle des matrices $\pi_{XYj}^{(k)}$, $k = 1, 2, \ldots, h$, même si l'on a « gagné un cran » par rapport à la proposition 7 en n'ayant à calculer $\pi_{XZ1}^{(k)}$ que pour $k \leq h - 1$ au lieu de $\pi_{XYj}^{(k)}$ pour $k \leq h$. C'est pourquoi on démontre que la propriété $Y \underset{(h)}{\not\longrightarrow} X$ peut aussi être caractérisée par la condition (iii) qui est plus intéressante, car elle ne fait intervenir que les coefficients autorégressifs initiaux, i.e. les blocs des matrices π_j de la représentation (3.2).

Le critère (iii) permet en outre de préciser le lien entre notre concept de « non-causalité à tout horizon » et le concept de non-causalité « à la Sims (1980) » en termes de coefficients de la forme moyenne mobile (« impulse response coefficients »). On considérera pour cela les séries formelles :

(3.11) $\pi(z) = I_n - \sum_{k=1}^{\infty} \pi_k z^k, \quad \psi(z) = \pi(z)^{-1} = I_n + \sum_{k=1}^{\infty} \psi_k z^k.$

Ces séries formelles, appliquées à l'opérateur retard, caractérisent la représentation autorégressive (3.2), $\pi(L)W(t) = a(t)$, et éventuellement une représentation moyenne mobile, $W(t) = \psi(L)a(t)$, lorsque la série $\sum_{k=1}^{\infty} \psi_k a(t-k)$ est bien définie comme limite en moyenne quadratique.

Il est, dans tous les cas, possible de partitionner ces séries formelles matricielles suivant les sous-vecteurs X(t), Y(t), Z(t) de W(t) et les sous-vecteurs associés $a_X(t)$, $a_Y(t)$, $a_Z(t)$ du processus d'innovation a(t). On obtient ainsi la représentation autorégressive

(3.12)
$$\pi_{XX}(L)\,X(t) + \pi_{XY}(L)\,Y(t) + \pi_{XZ}(L)\,Z(t) = a_X(t),$$
$$\pi_{YX}(L)\,X(t) + \pi_{YY}(L)\,Y(t) + \pi_{YZ}(L)\,Z(t) = a_Y(t),$$
$$\pi_{ZX}(L)\,X(t) + \pi_{ZY}(L)\,Y(t) + \pi_{ZZ}(L)\,Z(t) = a_Z(t),$$

et, éventuellement, une représentation moyenne mobile :

(3.13)
$$X(t) = \psi_{XX}(L)\,a_X(t) + \psi_{XY}(L)\,a_Y(t) + \psi_{XZ}(L)\,a_Z(t),$$
$$Y(t) = \psi_{YX}(L)\,a_X(t) + \psi_{YY}(L)\,a_Y(t) + \psi_{YZ}(L)\,a_Z(t),$$
$$Z(t) = \psi_{ZX}(L)\,a_X(t) + \psi_{ZY}(L)\,a_Y(t) + \psi_{ZZ}(L)\,a_Z(t).$$

Le critère usuel de non-causalité de Y vers X à l'horizon 1 est la nullité de la série $\pi_{XY}(L)$ qui assure que le passé de Y ne figure pas dans l'équation de X dans (3.12). Mais, d'après (3.12), on peut écrire formellement :

$$Z(t) = \pi_{ZZ}(L)^{-1}\,[a_Z(t) - \pi_{ZX}(L)\,X(t) - \pi_{ZY}(L)\,Y(t)]$$

où il est assuré que la série converge dès que la matrice $\pi_{ZZ}(z)$ est inversible pour $|z| \le 1$. En reportant dans l'équation de X(t), cela donne :

$$\left[\pi_{XX}(L) - \pi_{XZ}(L)\pi_{ZZ}(L)^{-1}\pi_{ZX}(L)\right]X(t) + \left[\pi_{XY}(L) - \pi_{XZ}(L)\pi_{ZZ}(L)^{-1}\pi_{ZY}(L)\right]Y(t)$$
$$= a_X(t) - \pi_{XZ}(L)\,\pi_{ZZ}(L)^{-1}\,a_Z(t).$$

On voit que, même si $\pi_{XY}(z) \equiv 0$, le passé de Y peut intervenir indirectement dans la détermination optimale de X dans le cas où la série formelle $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z)$ serait non nulle. Une telle démarche d'élimination de Z dans l'équation de X étant la première étape d'une inversion formelle de la série $\pi(z)$, on a donc l'intuition du résultat suivant (démontré de façon rigoureuse en annexe).

Proposition 9 (Conditions nécessaires de non-causalité à tout horizon) : Sous les hypothèses du théorème 1 et pourvu que la série $\pi(z)$ converge lorsque z est interprété comme un nombre complexe tel que $|z| < \delta$ (où δ est une constante positive), les trois conditions suivantes sont équivalentes et fournissent chacune une *condition nécessaire* pour que Y ne cause pas X à tout horizon :

(i) $\pi_{XY}(z) \equiv 0$ et $\pi_{XZ}(z)\,\pi_{ZZ}(z)^{-1}\,\pi_{ZY}(z) \equiv 0$,

(ii) $\pi_{XY}(z) \equiv 0$ et $\psi_{XY}(z) \equiv 0$,

(iii) $\psi_{XY}(z) \equiv 0$ et $\psi_{XZ}(z)\,\psi_{ZZ}(z)^{-1}\,\psi_{ZY}(z) \equiv 0$,

où la notation $\equiv$ signifie que les coefficients correspondants des séries formelles de chaque côté de $\equiv$ sont égaux.

Notons que la condition (iii) se déduit de (i) en permutant les rôles de π et ψ, ce qui est possible puisque, d'après (ii), π et ψ jouent un rôle symétrique. Ces conditions, qui éliminent à la fois l'effet *direct* de Y sur X (en annulant ses coefficients dans la représentation autorégressive) et ses effets *indirects* (en annulant les coefficients de ses innovations dans la représentation moyenne mobile), peuvent apparaître suffisantes pour que Y ne cause pas X à tout horizon.

Mais on va montrer, maintenant, grâce au théorème 1, que malgré les apparences, ces conditions ne sont pas en général suffisantes *dès que Z est multivarié,* du fait des éventuelles relations de causalité internes à Z. L'intérêt du critère (iii) du théorème 1 est précisément de fournir une condition de non-causalité à tout horizon qui est ***nécessaire et suffisante*** et qui, en outre, ne suppose pas l'existence d'une représentation moyenne mobile. Les matrices $R_{XZ}^{(k)}$, $k \in \mathbb{N}$, introduites par le théorème 1, peuvent en effet donner lieu à la définition d'une série formelle qu'il est aisé de relier à celles considérées dans la proposition 9.

Proposition 10 : Si on note

$$R_{XZ}(z) = \sum_{k=1}^{\infty} R_{XZ}^{(k)} z^k ,$$

où les matrices $R_{XZ}^{(k)}$ sont définies dans le théorème 1, la relation suivante est satisfaite :

$$R_{XZ}(z) \equiv -\pi_{XZ}(z)\, \pi_{ZZ}(z)^{-1} .$$

Par conséquent, la *condition nécessaire* de non-causalité de Y vers X à tout horizon énoncée dans la proposition 9 [voir en particulier sa forme (i)] s'écrit :

$$\text{(3.14)} \qquad \pi_{XYj} = 0 ,\ \forall j \geq 1 ,\ \text{et} \sum_{j=1}^{k-1} R_{XZ}^{(k-j)}\, \pi_{ZYj} = 0,\ \forall k > 1 .$$

Ceci n'est clairement pas, en général, une condition suffisante, puisque le critère (iii) du théorème 1 s'écrit :

$$\text{(3.15)} \qquad \pi_{XYj} = 0 ,\ \forall j \geq 1 ,\ \text{et}\ R_{XZ}^{(k)}\, \pi_{ZYj} = 0,\ \forall k \geq 1 ,\ \forall j \geq 1 .$$

La comparaison de (3.14) et (3.15) montre bien, en particulier, que la condition $\pi_{XY}(z) = \psi_{XY}(z) \equiv 0$ est en général insuffisante pour assurer que Y ne cause pas X à tout horizon.

Outre le cas trivial sans environnement (par de tierces variables Z) où les conditions $\pi_{XY}(z) \equiv 0$ et $\psi_{XY}(z) \equiv 0$ sont équivalentes et assurent la non-causalité à tout horizon, il y a cependant un cas important où (3.14) équivaut à (3.15) : celui où l'environnement est défini par un processus Z *univarié*. Dans ce cas, la condition de la proposition 9,

$$\pi_{XZ}(z)\, \pi_{ZZ}(z)^{-1}\, \pi_{ZY}(z) \equiv 0 ,$$

écrit, pour chaque composante x_i, $i = 1, \dots, m$, de X et chaque composante y_ℓ, $\ell = 1, \dots, p$, de Y, la nullité du produit de deux séries entières univariées, à savoir $\pi_{x_iZ}(z)\pi_{ZZ}(z)^{-1}$ et $\pi_{Zy_\ell}(z)$. Le produit de deux séries entières univariées ne pouvant être nul que si au moins l'une des deux est nulle (les zéros d'une fonction analytique sont isolés), on a donc alors nécessairement $R_{x_iZ}^{(k)} = 0$, $\forall k \geq 1$, ou encore $\pi_{Zy_\ell j} = 0$, $\forall j \geq 1$. Et, dans tous les cas,

$$R_{x_iZ}^{(k)}\, \pi_{Zy_\ell j} = 0 ,\ \forall k \geq 1 ,\ \forall j \geq 1 .$$

Ceci pouvant être fait pour tout $i = 1, 2, \dots, m$ et tout $\ell = 1, 2, \dots, p$, on en déduit que lorsque Z est univarié, la condition de la proposition 9 implique que

$$R_{XZ}^{(k)}\ \pi_{ZYj} = 0,\ \forall k \geq 1 ,\ \forall j \geq 1 ;$$

autrement dit, les conditions (3.14) et (3.15) sont dans ce cas équivalentes. On a ainsi démontré la proposition 11 suivante.

Proposition 11 (Caractérisations de la non-causalité à tout horizon pour Z univarié) : Sous les hypothèses du théorème 1, dans le cas d'un processus Z univarié et pourvu que la série $\pi(z)$ converge lorsque z est interprété comme un nombre complexe tel que $|z| < \delta$ (où δ est une constante positive), les quatre conditions suivantes sont équivalentes :

(i) Y ne cause pas X à tout horizon relativement à I_{XZ} ;

(ii) $\pi_{ZY}(z) \equiv 0$ et $\pi_{XZ}(z)\, \pi_{ZZ}(z)^{-1}\, \pi_{ZY}(z) \equiv 0$;

(iii) $\pi_{XY}(z) \equiv 0$ et $\psi_{XY}(z) \equiv 0$;

(iv) $\psi_{XY}(z) \equiv 0$ et $\psi_{XZ}(z)\, \psi_{ZZ}(z)^{-1}\, \psi_{ZY}(z) \equiv 0$.

Le cas où Z est multivarié est plus délicat car il implique que l'on prenne en compte les relations de causalité internes à Z. Ce problème sera étudié plus en détail dans la section 4 relative à la séparabilité.

Il convient, pour conclure cette section, que la condition nécessaire de la proposition 9, ainsi que le critère (iii) du théorème 1, fournissent des conditions empiriquement testables, car elles s'expriment en pratique par un nombre fini de contraintes sur les paramètres du processus considéré. Cela suppose, bien sûr, que l'on se place dans un modèle statistique défini lui-même à partir d'un nombre fini de paramètres, ce qui, en se limitant ici au contexte de processus stationnaires au second ordre, conduit à des représentations VAR(P), MA(Q) ou ARMA(P, Q).

La représentation autorégressive ayant été privilégiée dans le présent travail, on se contentera de détailler ici le cas d'un processus VAR(P). On vérifie d'abord aisément la proposition suivante.

Proposition 12 : Si $\{W(t) : t \in \mathbb{Z}, t > \omega\}$ est un processus autorégressif d'ordre P défini par (3.2) conformément aux hypothèses du théorème 1 (avec $\pi_k = 0$ pour $k > P$), les deux conditions suivantes sont équivalentes :

(i) Y ne cause pas X à tout horizon relativement à I_{XZ} ;

(ii) Y ne cause pas X jusqu'à l'horizon rP+1 relativement à I_{XZ}, où r désigne la dimension du processus Z(t) des variables auxiliaires.

Notons qu'un résultat de ce type aurait été trivial dans le contexte d'un processus moyenne mobile d'ordre Q ($\psi_k = 0$ pour $k > Q$), puisqu'il ne pourrait alors apparaître de causalité à un horizon supérieur à Q : pour $k > Q$, $W(t + k)$ est dans ce cas non corrélé avec tout élément de $I_W(t)$. Un processus autoregressif d'ordre P étant en général une moyenne mobile d'ordre infini, le résultat de la proposition 12 est moins évident (en dehors du cas limite $r = 0$ qui redonne la proposition 4). Il vient du fait que les conditions de non-causalité jusqu'à l'horizon h, portant sur les h-1 monômes de plus bas degré de $R_{XZ}(z) = -\pi_{XZ}(z)\,\pi_{ZZ}(z)^{-1}$ portent en fait sur les h-1 monômes de plus bas degré de $R^*_{XZ}(z) = -\pi_{XZ}(z)\,\pi^*_{ZZ}(z)$ où $\pi^*_{ZZ}(z)$ désigne la matrice adjointe de $\pi_{ZZ}(z)$. Comme les coefficients de la matrice $R^*_{XZ}(z)$ sont des polynômes de degré majoré par rP, on voit que la non-causalité à tout horizon est annexée dès qu'il y a non-causalité jusqu'à un horizon h avec $h-1 \geq rP$.

Il n'est pas surprenant d'obtenir un résultat analogue pour la condition nécessaire fournie par la proposition 9.

Proposition 13 : Sous les hypothèses de la proposition 12, la condition nécessaire de la proposition 9

$$\pi_{XY}(z) \equiv 0 \quad \text{et} \quad \pi_{XZ}(z)\,\pi_{ZZ}(z)^{-1}\,\pi_{ZY}(z) \equiv 0$$

est équivalente à l'ensemble fini de contraintes suivant :

(i) $\pi_{XYj} = 0$, pour $j = 1, 2, \dots, P$;

(ii) les coefficients de z^k dans $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z)$ sont nuls pour $k = 1, 2, \dots, rP+P$.

Ce résultat est à rapprocher de celui de Lütkepohl (1990) énoncé en termes des coefficients de la représentation moyenne mobile. Ainsi, la vérification empirique de la condition nécessaire de la proposition 9 implique au plus Pmp + (r P+P)mp = Pmp(r+2) contraintes à tester sur les coefficients du polynôme autorégressif. En revanche, la proposition 12 suggère que, pour tester que Y ne cause pas X à tout horizon, on peut être amené à tester jusqu'à Pmp + r PP mp = Pmp(rP+1) contraintes, puisque la vérification de la non-causalité jusqu'à l'horizon rP+1 nécessite en pratique de vérifier en particulier la nullité des rPP matrices $R^{(k)}_{XZ}\,\pi_{ZYj}$, $k = 1, 2, \dots, rP$ et $j = 1, 2, \dots, P$.

On trouve, comme cela était prévu [comparer (3.14) et (3.15)], *strictement plus de contraintes* à tester pour tester la non-causalité à tout horizon que pour tester la condition nécessaire fournie par la proposition 9, *à l'exception des cas particuliers* P = 1 ou encore P = 2 et r = 1. Ce dernier résultat n'est pas surprenant, puisque dans le cas P = 1, $\pi_1^{(h)} = \pi_1^h = \psi_h$ pour tout $h \geq 1$, si bien que les contraintes $\pi_{XY}(z) \equiv 0$ et $\psi_{XY}(z) \equiv 0$ fournies par la proposition 9 suffisent effectivement à caractériser la non-causalité à tout horizon.

Dans le cas général, la majoration proposée pourrait être affinée et on pourrait montrer qu'en pratique le test de la non-causalité de Y vers X à tout horizon impliquerait souvent strictement moins de Pmp(rP+1) contraintes. On sait, par exemple, que dans le cas r = 1, la non-causalité de Y vers X à tout horizon est assurée, quelle que soit la valeur de P, par la vérification de 3Pmp contraintes [et non pas (P+1)Pmp contraintes], puisque d'après la proposition 11, les contraintes fournies par la proposition 9 suffisent dans ce cas à caractériser la non-causalité à tout horizon.

4. SÉPARABILITÉ GÉNÉRALISÉE ET CONDITIONS DE RANG

En pratique, l'économètre qui souhaite vérifier que *la propriété de non-causalité de* Y *vers* X, qu'il a validé à l'horizon 1 par un test des conditions usuelles $\pi_{XYj} = 0$, $\forall j \geq 1$, *est robuste quand on élargit l'horizon*, peut tester les *hypothèses emboîtées* obtenues en ajoutant au fur et à mesure les contraintes :

$$R_{XZ}^{(k)} \pi_{ZYj} = 0 , \forall j \geq 1 ,$$

pour k = 1, puis pour k = 1 et 2, etc. Ces hypothèses sont des hypothèses sous forme implicite par rapport aux coefficients autorégressifs π_j et peuvent donc théoriquement être testées à partir d'une estimation de ces coefficients. Cependant, ces contraintes s'exprimant comme des produits, leur rang n'est pas constant, ce qui complique à la fois l'interprétation et le calcul du niveau des tests. C'est pourquoi il est utile de compléter le théorème 1 par des *conditions nécessaires* de non-causalité de Y vers X (jusqu'à un horizon h ou à tout horizon) qui soient plus simples à tester en pratique; un test statistique rejetant ces conditions permettra alors sans ambiguïté de rejeter l'hypothèse de non-causalité associée. L'étude de telles conditions nécessaires sera d'autant plus intéressante en pratique que leur interprétation économique sera claire. Or, on va montrer maintenant que ces conditions s'interprètent effectivement bien, au regard d'un *concept de séparabilité* généralisant celui considéré dans la proposition 5.

Le concept de séparabilité introduit par la proposition 5 fournissant seulement une condition *suffisante* de non-causalité de Y vers X à tout horizon, on se propose ici, pour énoncer des conditions *nécessaires* de non-causalité, de chercher à construire des fonctions linéaires, $Z_1(t)$ et $Z_2(t)$, de Z(t) telles que $I_{XZ}(t) = I_{XZ_1Z_2}(t)$ et qui satisfont une certaine condition de séparabilité. Pour simplifier, on supposera que $(Z_1(t)', Z_2(t)')'$ est un vecteur de $\mathbb{R}^r$ déduit de Z(t) par un changement de base orthonormée :

$$Z_1(t) = Q_1 Z(t) , \; Z_2(t) = Q_2 Z(t) , \tag{4.1}$$

où $Q = [Q_1', Q_2']'$ est une matrice orthogonale de taille r. On définit ainsi un nouveau processus $(X(t)', Y(t)', Z_1(t)', Z_2(t)')'$, $t \in \mathbb{Z}$, qui admet aussi une représentation autorégressive infinie de type (3.2) caractérisée par une suite de matrices carrées $\tilde{\pi}_k$, $k \in \mathbb{N}$ (de taille p+m+r) qui conduisent à définir, à partir d'équations analogues

à (3.6), (3.7), (3.8) et (3.9) des matrices $\tilde{\pi}_{XXj}^{(h)}$, $\tilde{\pi}_{XYj}^{(h)}$, $\tilde{\pi}_{XZ_1j}^{(h)}$, $\tilde{\pi}_{XZ_2j}^{(h)}$, $\tilde{\pi}_{YXj}^{(h)}$ ... , $j \in \mathbb{N}$, $h \in \mathbb{N}$. Les formules de passage entre les matrices π et $\tilde{\pi}$ s'obtiennent immédiatement par identification : si U et V désignent l'une des deux lettres X et Y, on a pour tous les entiers positifs j et h :

$$\begin{aligned}
&\tilde{\pi}_{UVj}^{(h)} = \pi_{UVj}^{(h)} , \; \tilde{\pi}_{UZ_1j}^{(h)} = \pi_{UZj}^{(h)} Q_1' , \; \tilde{\pi}_{UZ_2j}^{(h)} = \pi_{UZj}^{(h)} Q_2' , \\
&\tilde{\pi}_{Z_1Uj}^{(h)} = Q_1 \pi_{ZUj}^{(h)} , \; \tilde{\pi}_{Z_2Uj}^{(h)} = Q_2 \pi_{ZUj}^{(h)} , \; \tilde{\pi}_{Z_1Z_2j}^{(h)} = Q_1 \pi_{ZZj}^{(h)} Q_2' , \\
&\tilde{\pi}_{Z_1Z_1j}^{(h)} = Q_1 \pi_{ZZj}^{(h)} Q_1' , \; \tilde{\pi}_{Z_2Z_1j}^{(h)} = Q_2 \pi_{ZZj}^{(h)} Q_1' , \; \tilde{\pi}_{Z_2Z_2j}^{(h)} = Q_2 \pi_{ZZj}^{(h)} Q_2' .
\end{aligned} \tag{4.2}$$

Dans l'espoir d'obtenir une propriété *nécessaire* de séparabilité, on a intérêt à définir Z_1 de *taille maximale* parmi ceux qui vérifient : $Y \underset{1}{\not\longrightarrow} Z_1$ relativement à I_{XZ}. On souhaite donc définir une matrice Q_1 de taille $r_1 \times r$, de plein rang ligne, et d'un nombre de lignes maximal r_1 parmi les matrices qui vérifient

$$\tilde{\pi}_{Z_1Yj} = Q_1 \pi_{ZYj} = 0, \forall j \geq 1.$$

Chaque ligne de Q_1 doit donc être orthogonale au sous-espace E_{ZY} de $\mathbb{R}^r$ engendré par les vecteurs π_{Zy_ij}, i = 1, 2, ... , p, $j \geq 1$. La définition « maximale » de Q_1 correspond donc à $\mathrm{Im}(Q_1') = E_{ZY}^{\perp}$, où $E_{ZY}^{\perp}$ est le sous-espace de $\mathbb{R}^r$ orthogonal à E_{ZY}. Par conséquent, on supposera que les r_1 colonnes de Q_1' (c'est-à-dire les transposées des r_1 lignes de Q_1) constituent une base orthonormée de $E_{ZY}^{\perp}$; les r_2 colonnes de Q_2' constituant alors une base orthonormée de E_{ZY} (où $r_2 = r - r_1$).

Définition 3 : On appelle *canaux de transmission* de l'information $Y(\omega, t]$ vers X les composantes de

$$Z_2(t) = Q_2 Z(t) = (Z_{2i}(t) : i = 1, \ldots , r_2) ,$$

où les transposées des r_2 lignes de Q_2 constituent une base orthonormée de l'espace E_{ZY} engendré par les vecteurs π_{Zy_ij}, i = 1, 2, ... , p, $j \geq 1$. Ces canaux sont caractérisés par la conjonction des deux propriétés :

(i) $Y \underset{1}{\longrightarrow} Z_{2i}$ relativement à I_{XZ}, pour $i = 1, 2, \dots, r_2$,

(ii) $Y \underset{1}{\not\longrightarrow} Z_1$ relativement à I_{XZ},

où $Z_1(t) = Q_1 Z(t)$ et $[Q_1', Q_2']'$ est une matrice orthogonale de taille r.

La définition 3 repose bien sûr sur certaines conventions de notations : si $r_1 = r$ ($\pi_{ZYj} = 0$ pour tout $j \geq 1$), on note $Z = Z_1$ et on a $Y \underset{1}{\not\longrightarrow} Z$; si $r_1 = 0$, on note $Z = Z_2$ et on a $Y \underset{1}{\longrightarrow} z_i$ pour $i = 1, 2, \dots, r$. La propriété de causalité $Y \underset{1}{\longrightarrow} Z_{2i}$ désigne évidemment la négation de la propriété $Y \underset{1}{\not\longrightarrow} Z_{2i}$.

L'emprunt à la physique du vocable de « canal de transmission de l'information » [emprunt déjà fait, par exemple, dans un contexte voisin par Ahn et Reinsel (1988)] permet de traduire l'intuition suivante : comme chaque composante Z_{2i} de Z_2 est « causée » par Y, elle est susceptible de transmettre une causalité indirecte de Y vers X. De plus, Y ne causant pas Z_1, une telle causalité indirecte ne peut que passer par Z_2. Cette intuition est bien confirmée par le corollaire ci-dessous du théorème 1 qui montre que si $Y \underset{1}{\not\longrightarrow} X$, une causalité de Y vers X à un horizon supérieur ne peut qu'être attribuée à une causalité de Z_2 vers X.

Corollaire 1.1 : Sous les hypothèses du théorème 1, les deux propriétés suivantes sont équivalentes pour tout horizon $h > 1$:

(i) $Y \underset{(h)}{\not\longrightarrow} X$ relativement à I_{XZ} ;

(ii) $Y \underset{1}{\not\longrightarrow} X$ relativement à I_{XZ}, et $\bar{\pi}^{(k)}_{XZ_2 1} = 0$ pour $k = 1, 2, \dots, h-1$.

La propriété

$$\pi^{(k)}_{XZ1}\, \pi_{ZYj} = 0 \text{ , pour } k = 1, 2, \dots, h-1 \text{ , } j \geq 1 \text{ ,}$$

qui apparaît dans la caractérisation (ii) du théorème 1, est en effet équivalente à

$$\pi^{(k)}_{XZ1}\, Q_2' = 0 \text{ , pour } k = 1, 2, \dots, h-1 \text{ ,}$$

car les r_2 colonnes de Q_2' engendrent le même espace E_{ZY} que les vecteurs $\pi_{Zy_i j}$, $i = 1, 2, \dots, p$, $j \geq 1$. D'où le corollaire ci-dessus, puisque d'après (4.2),

$$\bar{\pi}^{(k)}_{XZ_2 1} = \pi^{(k)}_{XZ1}\, Q_2' \text{ .}$$

Ce corollaire signifie en particulier qu'une *condition nécessaire et suffisante pour que* Y *ne cause pas* X *à tout horizon* (sachant que $Y \underset{1}{\not\longrightarrow} X$) est que

$$\bar{\pi}^{(k)}_{XZ_2 1} = 0, \ \forall k \geq 1 \text{ .} \tag{4.3}$$

La condition (4.3) est une forme affaiblie d'une hypothèse de non-causalité de Z_2 vers X, puisqu'elle signifie que pour tout h, $P[X(t+h) \mid I(t) + Y(\omega, t]]$ ne dépend pas de $Z_2(t)$. Mais cette prévision peut dépendre, en revanche, de $Z_2(t-1)$, $Z_2(t-2)$, ... On peut se demander comment il est possible que cette forme de causalité de Z_2 vers X n'introduise pas une causalité indirecte de Y vers X passant par Z_2 (puisque, par construction, Y cause chacune des composantes de Z_2). La réponse à cette question est contenue dans le résultat suivant.

Proposition 14 : Sous les hypothèses du théorème 1, si $Y \underset{(h)}{\not\longrightarrow} X$ relativement à I_{XZ} et $h > 2$, alors

$$\bar{\pi}_{XZ_2 j} + \sum_{\ell=1}^{j-1} \bar{\pi}_{XZ_1 j-\ell}\, \bar{\pi}^{(\ell)}_{Z_1 Z_2 1} + \sum_{\ell=1}^{j-1} \bar{\pi}_{XZ_2 j-\ell}\, \bar{\pi}^{(\ell)}_{Z_2 Z_2 1} = 0 \text{ , pour } j = 2, 3, \dots, h-1 \text{ .}$$

La proposition 14 montre, en effet, que si $Z_2(t+1-j)$ intervient dans le calcul de la prévision optimale de $X(t+1)$ sachant $I(t) + Y(\omega, t]$, c'est-à-dire si $Z_2(t+1-j)$ a un « effet causal » sur $X(t+1)$, cette causalité est *compensée* (dans le cas où $Y \underset{(h)}{\not\longrightarrow} X$ avec $h > j$) par des chaînes de causalité en sens inverse représentables par les schémas suivants :

$$Z_2(t+1-j) \longrightarrow Z_1(t+1-j+\ell) \longrightarrow X(t+1)$$

ou

$$Z_2(t+1-j) \longrightarrow Z_2(t+1-j+\ell) \longrightarrow X(t+1) \text{ .}$$

La proposition suivante prouve même plus précisément que ce mécanisme de compensation est impossible s'il n'existe pas de chaînes de causalité du type $Z_2 \longrightarrow Z_1 \longrightarrow X$.

Proposition 15 : Supposons que les hypothèses du théorème 1 sont satisfaites. (i) Si $Y \underset{(h)}{\nrightarrow} X$ relativement a I_{XZ} et si $\sum_{\ell=1}^{j-1} \tilde{\pi}_{XZ_1 j-\ell}\, \tilde{\pi}^{(\ell)}_{Z_1 Z_2 1} = 0$ pour $j = 2, 3, \ldots, h-1$, alors $\tilde{\pi}_{XZ_2 j} = 0$ pour $j = 1, 2, \ldots, h-1$. (ii) *En particulier, si* Y *ne cause pas* X *à tout horizon*, au moins une des deux propriétés suivantes est vraie :

(P1) $Y \underset{1}{\nrightarrow} Z_1$ relativement à I_{XZ} et $Z_2 \underset{1}{\nrightarrow} X$ relativement à I_{XZ_1Y} (propriété de *séparabilité*);

(P2) $Y \underset{1}{\nrightarrow} Z_1$ relativement à I_{XZ} et il existe $j \geq 2$ tel que

$$\sum_{\ell=1}^{j-1} \tilde{\pi}_{XZ_1 j-\ell}\, \tilde{\pi}^{(\ell)}_{Z_1 Z_2 1} \neq 0 .$$

La propriété de séparabilité P1 est donc *nécessaire* dans les deux cas suivants :

(1) $Z_1 \underset{1}{\nrightarrow} X$ relativement à I_{XZ_2Y} (i.e. $\tilde{\pi}_{XZ_1 j} = 0$ pour $j \geq 1$);

(2) pour tout $\ell \geq 1$, $\pi^{(\ell)}_{ZZ1}$ est une matrice scalaire (c'est-à-dire proportionnelle à la matrice identité de taille r), puisqu'alors

$$\tilde{\pi}^{(\ell)}_{Z_1 Z_2 1} = Q_1\, \pi^{(\ell)}_{ZZ1}\, Q_2' = 0 .$$

Cette dernière condition est en particulier automatiquement vérifiée si Z est univarié, cas limite où Z ne peut qu'être égal à Z_1 ou Z_2.

Corollaire 15.1 : Si les hypothèses du théorème 1 sont satisfaites et si Z est univarié ($r = 1$), alors Y ne cause pas X à tout horizon *si et seulement si* au moins une des deux propositions suivantes est vérifiée :

(i) $Y \underset{1}{\nrightarrow} \begin{bmatrix} X \\ Z \end{bmatrix}$ relativement à I_{XZ} ;

(ii) $\begin{bmatrix} Y \\ Z \end{bmatrix} \underset{1}{\nrightarrow} X$ relativement à I_X .

Il convient de remarquer que les conditions (i) et (ii) ci-dessus sont chacune (comme cas extrêmes de la condition de séparabilité définie dans la proposition 5) des *conditions suffisantes* pour que $Y \underset{\infty}{\nrightarrow} X$ relativement à I_{XZ}, quelque soit la dimension de Z. Lorsque Z est univarié, on peut affirmer que *nécessairement*, au moins une des deux conditions est vérifiée (si Y ne cause jamais X), parce que dans ce cas particulier, les *compensations internes à* Z ne peuvent se réaliser. Il est intéressant de noter ici que la vérification d'au moins une des conditions (i) et (ii) correspond à la définition de la « non-causalité » proposée par Hsiao (1982, p. 247, définition 3) dans le contexte où les processus X, Y, Z sont univariés. Mais cette définition apparaît trop restrictive dans le cas où Z est multivarié, puisqu'il est alors possible qu'elle ne soit pas vérifiée alors que pourtant Y ne cause pas X *à tout horizon*.

Lorsque Z est multivarié, on peut vouloir, dans un premier temps, démêler l'écheveau des canaux de transmission de l'information de Y vers X à partir de *tests de rang* sur les matrices qui définissent la représentation autorégressive [voir, par exemple, Ahn et Reinsel (1988)]. Le théorème 1 fournit effectivement un ensemble de *conditions de rang nécessaires* pour que Y ne cause pas X jusqu'à l'horizon h. Il suffit, pour les obtenir, de considérer les sous-espaces suivants de $\mathbb{R}^r$. Outre l'espace E_{ZY} engendré par les colonnes des matrices π_{ZYj}, $j \geq 1$, on considère aussi pour $h \geq 2$: d'une part, l'espace $E^{(h-1)}_{XZ}$ engendré par les transposées des lignes $\pi^{(k)}_{x_i Z1}$, $i = 1, 2, \ldots, m$, $k = 1, 2, \ldots, h-1$, des matrices $\pi^{(k)}_{XZ1}$, $1 \leq k \leq h-1$; d'autre part, l'espace $F^{(h-1)}_{XZ}$ engendré par les transposées des lignes $R^{(k)}_{x_i Z}$, $i = 1, 2, \ldots, m$, $k = 1, 2, \ldots, h-1$, des matrices $R^{(k)}_{XZ}$, $1 \leq k \leq h-1$.

D'après le théorème 1, si $Y \underset{1}{\nrightarrow} X$, on a : $Y \underset{(h)}{\nrightarrow} X$ relativement à I_{XZ} $\Leftrightarrow$ $E^{(h-1)}_{XZ} \perp E_{ZY} \Leftrightarrow F^{(h-1)}_{XZ} \perp E_{ZY}$. L'orthogonalité de ces sous-espaces de $\mathbb{R}^r$ implique les conditions de dimension :

$$(4.4) \qquad \dim(E_{XZ}^{(h-1)}) + \dim(E_{ZY}) \leq r \,, \; \dim(F_{XZ}^{(h-1)}) + \dim(E_{ZY}) \leq r \,.$$

Ces conditions de dimension peuvent, bien sûr, être facilement traduites en termes de conditions de rang en remarquant que $\dim(E_{XZ}^{(h-1)}) = \text{rang}(P_{XZ}^{(h-1)})$ et $\dim(F_{XZ}^{(h-1)}) = \text{rang}(Q_{XZ}^{(h-1)})$, où les matrices $P_{XZ}^{(h-1)}$ et $Q_{XZ}^{(h-1)}$ sont définies par

$$(4.5) \qquad P_{XZ}^{(h-1)} = \begin{bmatrix} \pi_{XZ1}^{(1)} \\ \pi_{XZ1}^{(2)} \\ \vdots \\ \pi_{XZ1}^{(h-1)} \end{bmatrix}, \; Q_{XZ}^{(h-1)} = \begin{bmatrix} R_{XZ}^{(1)} \\ R_{XZ}^{(2)} \\ \vdots \\ R_{XZ}^{(h-1)} \end{bmatrix}.$$

En outre, pour tout $k \geq 1$, $\dim(E_{ZY})$ est minorée par le rang de la matrice

$$(4.6) \qquad S_{ZY}^{(k)} = [\pi_{ZY1}, \pi_{ZY2}, \dots, \pi_{ZYk}] \,.$$

Il résulte donc immédiatement des conditions (4.4) la proposition suivante.

Proposition 16 (Conditions de rang pour la non-causalité jusqu'à un horizon h) : Sous les hypothèses du théorème 1, chacune des conditions suivantes implique le rejet de l'hypothèse ($Y \underset{(h)}{\not\longrightarrow} X$ relativement à I_{XZ}) de non-causalité jusqu'à l'horizon h ($h \geq 2$) :

(i) $\dim(E_{XZ}^{(h-1)}) + \dim(E_{ZY}) > r$;

(ii) $\dim(F_{XZ}^{(h-1)}) + \dim(E_{ZY}) > r$;

(iii) $\text{rang}(P_{XZ}^{(h-1)}) + \dim(E_{ZY}) > r$;

(iv) $\text{rang}(P_{XZ}^{(h-1)}) + \text{rang}(S_{ZY}^{(k)}) > r$, où $k \geq 1$;

(v) $\text{rang}(Q_{XZ}^{(h-1)}) + \dim(E_{ZY}) > r$;

(vi) $\text{rang}(Q_{XZ}^{(h-1)}) + \text{rang}(S_{ZY}^{(k)}) > r$, où $k \geq 1$.

5. NON-CAUSALITÉS DANS UN PROCESSUS ARIMA

La modélisation autorégressive étant, comme on l'a montré dans les sections 3 et 4, bien adaptée à l'étude des propriétés de non-causalité linéaire, il est utile d'étudier plus en détail le cas de processus ARIMA, qui ne sont pas nécessairement stationnaires, en particulier du fait de la présence de racines unités. Ceci conduit à considérer, de façon plus explicite, des représentations autorégressives *conditionnelles à des valeurs initiales*, la suite des coefficients autorégressifs n'étant pas nécessairement absolument sommable.

Définition 4 : On dira que le processus centré $\{W(t) : t \in \mathbb{N}_0\}$ de dimension n admet une *représentation linéaire* par rapport à un bruit a(t) si on peut écrire, pour tout entier $t \geq 0$:

$$(5.1) \qquad W(t) = a(t) + \sum_{j=1}^{t} \pi_j \, W(t-j) + C(t)\, \overline{W}(-1) \,,$$

où π_j et C(t) sont des matrices déterministes ($j \in \mathbb{N}$, $t \in \mathbb{N}_0$), $\{a(t) : t \in \mathbb{N}_0\}$ est une suite de vecteurs aléatoires dont les composantes appartiennent à L^2, de moyenne nulle et non corrélés entre eux, $\overline{W}(-1)$ est un vecteur aléatoire appelé « *condition initiale* » non corrélé avec les « valeurs futures » du bruit a(t), $t \in \mathbb{N}_0$.

Dans la précédente définition et dans la suite de ce texte, nous adoptons la convention que toute sommation de la forme $\sum_{k=s}^{t}$ est nulle lorsque $t < s$. Il est clair que la représentation (5.1) permet de caractériser le processus W(t) comme fonction linéaire des valeurs présentes et passées du bruit et des conditions initiales. C'est l'objet de la proposition 17 ci-dessous.

Proposition 17 : (i) Si le processus $\{W(t), t \in \mathbb{N}_0\}$ vérifie (5.1), on a pour tout $t \geq 0$:

$$(5.2) \qquad W(t) = a(t) + \sum_{j=1}^{t} H_j \, a(t-j) + B(t)\, \overline{W}(-1) \,,$$

où H_j et B(t) sont des matrices déterministes définies par

(5.3) $$H_1 = \pi_1\ ,\quad B(0) = C(0)\ ,$$
$$H_j = \pi_j + \sum_{i=1}^{j-1} \pi_{j-i} H_i\ ,\ \text{pour } j \geq 2\ ,$$
$$B(t) = \sum_{j=1}^{t} \pi_j\, B(t-j) + C(t)\ ,\ \text{pour } t \geq 1\ .$$

(ii) Réciproquement, un processus centré $\{W(t),\ t \in \mathbb{N}_0\}$ défini par une équation de type (5.2) admet aussi une représentation linéaire de type (5.1) avec

(5.4) $$\pi_1 = H_1\ ,\quad C(0) = B(0)\ ,$$
$$\pi_j = H_j - \sum_{i=1}^{j-1} H_{j-i}\, \pi_i\ ,\ \text{pour } j \geq 2\ ,$$
$$C(t) = B(t) - \sum_{j=1}^{t} H_j C(t-j)\ ,\ \text{pour } t \geq 1\ .$$

On considérera, dans toute cette section, un processus $W(t) = (X(t)', Y(t)', Z(t)')'$, $t \in \mathbb{N}_0$, défini par (5.1) (ou de manière équivalente par (5.2)). La double représentation (5.1) et (5.2) permet de définir sans ambiguïté une notion « linéaire » d'information disponible à la date t.

Soit $X[0, t]$ l'espace de Hilbert engendré par les composantes de $X(\tau)$, $\tau = 0, 1, \dots, t$, $Y[0, t]$ celui engendré par les composantes de $Y(\tau)$, $\tau = 0, 1, \dots, t$, et $\overline{W}[-1]$ celui déterminé par les composantes de $\overline{W}(-1)$. L'information de référence sera ici

(5.5) $$I(t) = X[0, t] + Z[0, t] + \overline{W}[-1]\ ,$$

i.e. l'espace de Hilbert engendré par les composantes de $(X(\tau)', Z(\tau)')'$, $\tau = 0, 1, 2, \dots, t$ et celles de $\overline{W}(-1)$. La conjonction de (5.1) et (5.2) montre que $I(t) + Y[0, t]$ est aussi l'espace vectoriel engendré par les composantes de $a(\tau)$, $\tau = 0, 1, \dots, t$ et celles de $\overline{W}(-1)$. Comme ces composantes sont, par hypothèse, non corrélées avec $a(t + 1)$, on a donc :

$$P[a(t+1) \mid I(t) + Y[0, t]] = 0\ .$$

En appliquant l'opérateur $P[\ .\ \mid I(t) + Y[0, t]]$ à l'équation (5.1), on en déduit la proposition 18 ci-dessous.

Proposition 18 : Si le processus $\{W(t) : t \in \mathbb{N}_0\}$ vérifie (5.1) et $W(t) = (X(t)', Y(t)', Z(t)')'$, on a pour tout $t \geq 0$:

$$P[W(t+1) \mid I(t) + Y[0, t]] = \sum_{j=1}^{t+1} \pi_j\, W(t+1-j) + C(t+1)\, \overline{W}(-1)\ .$$

où $I(t) = X[0, t] + Z[0, t] + \overline{W}[-1]$. Autrement dit, $\{a(t),\ t \in \mathbb{N}_0\}$, est le processus d'innovation de $W(t)$.

Outre qu'il se prête bien au calcul des prévisions linéaires, le cadre des représentations linéaires du type (5.1) est très général, puisqu'en particulier, tout processus (stationnaire ou non) admettant une représentation ARMA(p, q) est de ce type [voir Gouriéroux et Monfort (1990, pages 320-321), pour la démonstration de la proposition 19 ci-dessous].

Proposition 19 : Soit $\{W(t) : t \in \mathbb{N}_0\}$ un processus de dimension n admettant une représentation ARMA(p, q) :

$$W(t) - \sum_{j=1}^{p} \phi_j\, W(t-j) = a(t) - \sum_{i=1}^{q} \Theta_i\, a(t-i),\ t \geq 0\ ,$$

où les variables $W(-1), W(-2), \dots, W(-p)$ et $a(-1), a(-2), \dots, a(-q)$ sont non corrélées avec les valeurs futures $a(t)$, $t \geq 0$, et $\{a(t) : t \in \mathbb{N}_0\}$ est une suite de vecteurs aléatoires non corrélés entre eux. Alors $\{W(t) : t \in \mathbb{N}_0\}$ admet une représentation linéaire du type (5.1) par rapport à $a(t)$ avec $\overline{W}(-1) = (W(-1)', W(-2)', \dots, W(-p)', a(-1)', \dots, a(-q)')'$.

Notons que la classe des processus ARMA(p, q), au sens de la proposition 19, contient en particulier tous les processus dits ARIMA parce que : d'une part, $\det \Theta(z) = \det [I_n - \sum_{i=1}^{q} \Theta_i z^i]$ a toutes ses racines à l'extérieur du disque unité; et d'autre part, $\det \phi(z) = \det [I_n - \sum_{j=1}^{p} \phi_j z^j]$ a toutes ses racines à l'extérieur du disque unité, sauf certaines égales à 1. Parmi ces processus, certains (appelés processus

ARIMA(p−d, d, q)) admettent une représentation ARMA stationnaire inversible après différenciation (quand la racine unité n'apparaît que par un facteur $(1 - z)^d$ dans $\phi(z)$) alors que, pour d'autres, il existe des relations de coïntégration.

Nous commencerons par donner des caractérisations des propriétés de non-causalité valables très généralement, *pour tout processus admettant une représentation linéaire de type (5.1)*, avant de mettre en évidence les propriétés spécifiques des processus ARIMA(p, d, q). On remarque d'abord que la représentation (5.1) permet un calcul des prévisions à l'horizon h analogue à celui caractérisé par la proposition 6, l'unique différence étant la prise en compte des conditions initiales. C'est l'objet de la proposition 20 ci-dessous.

Proposition 20 : Si le processus $\{W(t) : t \in \mathbb{N}_0\}$ vérifie (5.1) et $W(t) = (X(t)', Y(t)', Z(t)')'$, alors

$$(5.6) \qquad P[W(t+h) \mid I(t) + Y[0, t]] = \sum_{j=1}^{t+1} \pi_j^{(h)} W(t+1-j) + C^{(h)}(t+1)\,\overline{W}(-1),$$

pour tout $t \in \mathbb{N}_0$ et $h \in \mathbb{N}$, où $I(t) = X[0, t] + Z[0, t] + \overline{W}[-1]$ tel que défini pour (5.5) et où, pour tout $j \in \mathbb{N}$, les suites de matrices $\pi_j^{(h)}$ et $C^{(h)}(t+1)$, $h \in \mathbb{N}$, sont définies par récurrence :

$$(5.7) \qquad \pi_j^{(1)} = \pi_j\,,\ C^{(1)}(t+1) = C(t+1)\,,\ \pi_j^{(h+1)} = \pi_{j+h} + \sum_{\ell=1}^{h} \pi_{h-\ell+1}\,\pi_j^{(\ell)}\,,$$

$$C^{(h+1)}(t+1) = C(t+1+h) + \sum_{\ell=1}^{h} \pi_{h-\ell+1}\,C^{(\ell)}(t+1)\,.$$

Ces suites vérifient en outre les relations de récurrence

$$(5.8a) \qquad \pi_j^{(h+1)} = \pi_{j+1}^{(h)} + \pi_1^{(h)}\pi_j\,,\ h = 1, 2, \ldots,$$

$$(5.8b) \qquad C^{(h+1)}(t+1)\overline{W}(-1) = C^{(h)}(t+2)\overline{W}(-1) + \pi_1^{(h)}C(t+1)\,\overline{W}(-1)\,,\ h = 1, 2, \ldots.$$

On en déduit immédiatement un résultat analogue à la proposition 7 pour caractériser la non-causalité à l'horizon h.

Proposition 21 : Si le processus $\{W(t) : t \in \mathbb{N}_0\}$ vérifie (5.1), la condition

$$(5.9) \qquad \pi_{XYj}^{(h)} = 0,\ j = 1, 2, \ldots$$

est suffisante pour que $Y \underset{h}{\not\to} X$ relativement à I, où $I(t) = X[0, t] + Z[0, t] + \overline{W}[-1]$ tel que défini pour (5.5). Si, de plus, la matrice covariance $E[a(t)a(t)']$ est non singulière pour tout $t \geq 0$, la condition (5.9) est nécessaire et suffisante pour que $Y \underset{h}{\not\to} X$ relativement à I.

Les propositions 8, 9, 10, 11, le théorème 1 ainsi que leurs corollaires ne développant que des conséquences algébriques de la caractérisation (3.10) de la propriété $Y \underset{h}{\not\to} X$, tous ces résultats peuvent donc être repris pour caractériser la non-causalité dans le cadre (5.1). Il est important de noter, cependant, que l'introduction de *conditions initiales* affecte sensiblement le contenu sémantique de la propriété de non-causalité. La propriété « $Y \underset{h}{\not\to} X$ relativement à I » définie, conformément à (2.3), par

$$P[X(t+h) \mid I(t) + Y[0, t]] = P[X(t+h) \mid I(t)]$$

n'implique pas, en effet, que *les valeurs passées de* Y n'interviennent pas dans le calcul de la prévision (linéaire) optimale de X à l'horizon h. Plus précisément, si la nullité de $\pi_{XYj}^{(h)}$ pour tout j assure qu'effectivement $P[X(t+h) \mid I(t) + Y[0, t]]$ ne dépend pas des valeurs $Y(t+1-j)$ pour $j = 1, 2, \ldots, t+1$, cette prévision peut dépendre, en revanche, des valeurs passées de Y par l'intermédiaire des termes $C^{(h)}(t+1)\overline{W}(-1)$ relatifs aux conditions initiales. La proposition 19 a montré, par exemple, que $\overline{W}(-1)$ pouvait contenir en particulier l'information relative aux valeurs de $Y(-1), Y(-2), \ldots, Y(-p)$.

En ce sens, la propriété de non-causalité caractérisée par (5.9) ne porte que sur les *innovations* des processus X, Y, Z *à partir de la date* 0, mais pas sur les relations causales globales entre ces processus. Cette carence sémantique est due, en fait, à une *définition trop imprécise du rôle des conditions initiales*. Nous allons maintenant montrer que ce problème ne se pose pas si on spécifie davantage la représentation (5.1) en se plaçant néanmoins toujours dans un cadre très général, puisqu'on vérifiera qu'il contient en particulier l'ensemble des processus ARIMA(p, d, q).

Définition 6 : On dira que le processus centré $\{W(t) : t \in \mathbb{N}_0\}$ de dimension n admet une ***représentation linéaire intégrée d'ordre d*** (où $d \in \mathbb{N}$) par rapport à un bruit a(t) si on peut écrire, pour tout entier $t \geq 0$,

$$W(t) = a(t) + \sum_{j=1}^{t} \pi_j W(t-j) + C_d(t)\, \bar{W}_d(-1)$$

où :

(i) $\bar{W}_d(-1) = (W(-1)', W(-2)', \ldots, W(-d)')'$;

(ii) $\{a(t) : t \in \mathbb{N}_0\}$ est une suite de vecteurs aléatoires dont les composantes appartiennent à L^2, de moyenne nulle, non corrélés entre eux et avec $\bar{W}(-1)$;

(iii) la suite matricielle π_j, $j \in \mathbb{N}$, est obtenue par d différenciations successives d'une suite matricielle π_j^*, $j \in \mathbb{N}_0$, i.e. $\pi_j = \nabla^d \pi_j^*$, où $\nabla\pi_j^* = \pi_j^* - \pi_{j-1}^*$, avec pour $k = 0, 1, 2, \ldots, d-1$, la convention $\nabla^k \pi_0^* = -I_n$, matrice opposée de la matrice identité de taille n;

(iv) $C_d(t)$ est la matrice de dimension $n \times (nd)$ telle que

$$C_d(t)\, \bar{W}_d(-1) = -\sum_{\ell=0}^{d-1} (\nabla^{\ell}\pi_{\ell}^*)\, \nabla^{d-\ell-1} W(-1)$$

où $\nabla W(t)$, $t \in \mathbb{N}_0$, est le processus différencié $\nabla W(t) = W(t) - W(t-1)$.

La représentation linéaire intégrée d'ordre d permet en particulier de modéliser un processus W(t) dont le processus d fois différencié $\nabla^d W(t)$ est asymptotiquement équivalent à un processus stationnaire au second ordre régulier dont la décomposition de Wold est inversible, la spécification complète de W(t) étant obtenue par un choix naturel de conditions initiales. Considérons, en effet, un processus $\tilde{W}(t)$ dont la différence d'ordre d, $\nabla^d \tilde{W}(t)$, est un processus stationnaire du second ordre régulier d'innovation a(t) et dont la décomposition de Wold

(5.10) $$\nabla^d \tilde{W}(t) = a(t) + \sum_{j=1}^{\infty} H_j^* a(t-j)$$

est inversible, ce qui permet d'écrire une représentation autorégressive infinie

(5.11) $$\nabla^d \tilde{W}(t) = a(t) + \sum_{j=1}^{\infty} \pi_j^* \nabla^d \tilde{W}(t-j) .$$

On supposera, en outre, que les inversions (passage de (5.10) à (5.11) et réciproquement) sont obtenues par la méthode usuelle de division selon les puissances croissantes à partir de l'hypothèse de sommabilité absolue des séries :

(5.12) $$\sum_{j=1}^{\infty} \| \pi_j^* \| < +\infty \quad \text{et} \quad \sum_{j=1}^{\infty} \| H_j^* \| < +\infty .$$

Ce cadre recouvre bien, en particulier, celui des processus ARIMA(p, d, q), puisque si l'on a, conformément à la proposition 19,

(5.13) $$\phi(L)\, \tilde{W}(t) = \Theta(L) a(t) , \quad \phi(L) = (1-L)^d\, \psi(L)$$

où les polynômes $\psi(L)$ et $\Theta(L)$ ont toutes leurs racines à l'extérieur du disque unité, la représentation (5.10) (resp. (5.11)) peut être déduite de (5.13) en inversant $\psi(L)$ (resp. $\Theta(L)$).

Les équations (5.10) ou (5.11) ne définissent pas le processus $\tilde{W}$ de façon unique (l'opérateur ∇^d n'étant pas inversible); les solutions $\tilde{W}$ de (5.10) et (5.11) peuvent cependant être bien caractérisées à partir du résultat suivant.

Proposition 22 : Si un processus $\{\tilde{W}(t) : t \in \mathbb{Z}\}$ de dimension n vérifie (5.11), alors pour $h = 0, 1, 2, \ldots, d$, et pour tout $t \geq d$,

$$\nabla^{d-h}\tilde{W}(t) = a(t) + \sum_{j=1}^{t} (\nabla^h \pi_j^*)\, \nabla^{d-h}\tilde{W}(t-j) + C_d(t,h)\tilde{W}_d(-1) + \sum_{j=t+1}^{\infty} \pi_j^* \nabla^d \tilde{W}(t-j)$$

où $d \in \mathbb{N}_0$, $\tilde{W}_d(-1) = (\tilde{W}(-1)', \tilde{W}(-2)', \ldots, \tilde{W}(-d)')'$ et $C_d(t, h)$ est la matrice de dimension $n \times (nd)$ telle que

$$C_d(t,h)\tilde{W}_d(-1) = -\sum_{\ell=0}^{h-1} (\nabla^{\ell}\pi_{\ell}^*)\, \nabla^{d-\ell-1}\tilde{W}(-1) ,$$

en convenant que $C_d(t, 0)\tilde{W}_d(-1) = 0$, $\nabla\pi^*_j = \pi^*_j - \pi^*_{j-1}$, $\nabla^{k+1}\pi^*_j = \nabla^k\pi^*_j - \nabla^k\pi^*_{j-1}$ et $\nabla^k\pi^*_0 = -I_n$ pour $k = 0, 1, \dots, d-1$.

En prenant $h = d$, il résulte de la proposition 17 qu'un processus $\tilde{W}(t)$ vérifiant (5.11) vérifie aussi, pour $t \geq d$,

$$\tilde{W}(t) = a(t) + \sum_{j=1}^{t} (\nabla^d\pi^*_j)\, \tilde{W}(t-j) + C_d(t)\tilde{W}_d(-1) + \sum_{j=t+1}^{\infty} \pi^*_j\, \nabla^d\tilde{W}(t-j)\,, \tag{5.14}$$

où on définit $C_d(t) = C_d(t, d)$. La représentation linéaire intégrée d'ordre d de la définition 6,

$$W(t) = a(t) + \sum_{j=1}^{t} \pi_j\, W(t-j) + C_d(t)\, \tilde{W}_d(-1)\,, \tag{5.15}$$

avec $\tilde{W}_d(-1) = (\tilde{W}(-1)', \tilde{W}(-2)', \dots, \tilde{W}(-d)')'$ et $\pi_j = \nabla^d\pi^*_j$, peut donc être réinterprétée à partir de la représentation générale (5.11) en considérant que celle-ci a été *complétée par des conditions initiales* du type :

$$\nabla^d\, W(\tau) = 0 \text{ pour } \tau < 0\,.$$

Le choix de ces conditions initiales est bien naturel, puisqu'il s'agit de traduire l'idée que $\nabla^d\, W(t)$ est « approximativement » un processus stationnaire centré. D'autre part, d'un point de vue *asymptotique* (on considère les processus à une date t « grande »), ce choix est sans conséquences, puisque la comparaison de (5.15) et (5.14) montre que W(t) est *asymptotiquement équivalent à tout processus* $\tilde{W}(t)$ vérifiant (5.14) (pour $W_d(-1) = \tilde{W}_d(-1)$). En effet, $\sum_{j=t+1}^{\infty} \pi^*_j\nabla^d\tilde{W}(t-j)$ converge en moyenne quadratique vers zéro (quand t tend vers l'infini), car $\nabla^d\, \tilde{W}(t)$ est stationnaire et $\sum_{j=t+1}^{\infty} \| \pi^*_j \|$ converge vers zéro (reste d'une série convergente). Notons enfin que dans le cas particulier où les matrices π^*_j s'annulent à partir d'un certain ordre (cas des processus ARIMA(p, d, 0)), W et $\tilde{W}$ coïncident exactement à partir d'une certaine date (et ne sont pas seulement asymptotiquement équivalents).

Le cadre des processus admettant une représentation linéaire intégrée d'ordre d est particulièrement bien adapté à l'étude des propriétés de non-causalité. Notre résultat essentiel est, à ce sujet, le théorème 2 ci-dessous. L'argument central de la démonstration de ce dernier repose sur trois lemmes purement algébriques que nous énonçons d'abord comme propositions 23, 24 et 25, parce qu'ils paraissent importants pour l'étude des processus ARIMA, indépendamment de la problématique particulière du théorème 2.

Proposition 23 : Si, pour chaque entier positif j, on définit par la récurrence (3.7) une suite de matrices $\pi^{(h)}_j$, $h \in \mathbb{N}$, de dimension $n \times n$, cette dernière satisfait la relation :

$$\pi^{(k)}_j - \pi^{(k)}_{j-1} = (\nabla\pi_j)^{(k)} - (\nabla\pi_j)^{(k-1)}\,, \ \forall j\,, k \in \mathbb{N}\,, \tag{5.16}$$

où, pour chaque $j \geq 1$, la suite $(\nabla\pi_j)^{(k)}$, $k \geq 1$, est définie par la récurrence

$$(\nabla\pi_j)^{(1)} = \nabla\pi_j = \pi_j - \pi_{j-1}\,, \quad (\nabla\pi_j)^{(k+1)} = \nabla\pi_{j+k} + \sum_{\ell=1}^{k} (\nabla\pi_{k-\ell+1})\,(\nabla\pi_j)^{(\ell)}\,,$$

en convenant de noter $\pi^{(k)}_0 = 0$ pour $k > 1$, $\pi_0 = \pi^{(1)}_0 = -I_n$ et $(\nabla\pi_j)^{(0)} = 0$.

Proposition 24 : Sous les hypothèses de la proposition 23, les deux propriétés suivantes sont équivalentes :

(i) $\pi^{(k)}_{XYj} = 0$, pour $j \in \mathbb{N}$ et $k = 1, 2, \dots, h$,

(ii) $(\nabla\pi_j)^{(k)}_{XY} = 0$, pour $j \in \mathbb{N}$ et $k = 1, 2, \dots, h$,

où les matrices $\pi^{(k)}_{XYj}$ et $(\nabla\pi_j)^{(k)}_{XY}$ sont obtenues en partitionnant $\pi^{(k)}_j$ et $(\nabla\pi_j)^{(k)}$ de façon analogue à $\pi^{(h)}_j$ en (3.9).

Proposition 25 : Soit π_t, $t \in \mathbb{N}$, une suite de matrices $n \times n$ obtenues par d différentiations successives d'une autre suite matricielle π_t^*, $t \in \mathbb{N}_0$, i.e. $\pi_t = \nabla^d \pi_t^*$, où $\nabla \pi_t^* = \pi_t^* - \pi_{t-1}^*$, avec pour $k = 0, 1, \dots, d-1$, la convention $\nabla^k \pi_0^* = -I_n$. Si

$$\pi_{XYt}^{*(k)} = 0, \text{ pour } k = 1, 2, \dots, h, \ \forall t \in \mathbb{N},$$

alors, pour $k = 1, 2, \dots, h-1$ et $\ell = 0, 1, \dots, d$,

$$\pi_{XZ1}^{(k)} (\nabla^\ell \pi_t^*)_{ZY} = 0, \ \forall t \in \mathbb{N},$$

où les matrices $\pi_{XYt}^{*(k)}$, $\pi_{XZ1}^{(k)}$ et $(\nabla^\ell \pi_t^*)_{ZY}$ sont obtenues par une partition analogue à celle figurant en (3.9).

Nous pouvons maintenant démontrer le théorème 2.

Théorème 2 : Soit $\{W(t) : t \in \mathbb{N}_0\}$ un processus défini conformément à la définition 6 à partir d'un bruit blanc $a(t)$ de *variance régulière*, avec $W(t) = (X(t)', Y(t)', Z(t)')'$, et soit $I(t) = X[0, t] + Z[0, t] + \overline{W}[-1]$ tel que défini pour (5.5) avec $\overline{W}(-1) = \overline{W}_d(-1)$. Alors, les quatre propriétés suivantes sont équivalentes :

(i) $Y \underset{(h)}{\not\rightarrow} X$ relativement à I ;

(ii) pour tout $t \in \mathbb{N}_0$ et pour $k = 1, 2, \dots, h$,

$$P[X(t+k) \mid I(t) + Y[0, t]] = P[X(t+k) \mid X[-d, t] + Z[-d, t]] ;$$

(iii) $\pi_{XYj}^{*(k)} = 0$ pour $k = 1, 2, \dots, h$ et pour tout $j \geq 1$;

(iv) pour tout $t \in \mathbb{N}_0$, les valeurs de $Y(\tau)$, $-d \leq \tau \leq t$, n'interviennent pas dans le calcul des prévisions optimales $P[X(t+k) \mid I(t) + Y[0, t]]$, $k = 1, 2, \dots, h$.

Les caractérisations (ii) et (iv) de la propriété $Y \underset{(h)}{\not\rightarrow} X$ résolvent bien le problème d'interprétation de cette propriété que nous avions souligné à la suite de la proposition 16. Si $Y \underset{(h)}{\not\rightarrow} X$, les valeurs passées de Y, *y compris les conditions initiales sur* Y, n'interviennent pas dans le calcul des prévisions optimales de X aux horizons $k = 1, 2, \dots, h$.

La caractérisation (iii) (avec des notations $\pi_{XYj}^{*(k)}$ analogues aux notations $\pi_{XYj}^{(k)}$ utilisées dans le reste de l'article) montre en outre que la propriété de non-causalité jusqu'à l'horizon h, définie à partir des séries brutes X, Y et Z, s'interprète aussi en termes de non-causalité sur les séries différenciées $\nabla^d X$, $\nabla^d Y$ et $\nabla^d Z$. Dans le cas où celles-ci sont (asymptotiquement) stationnaires, la condition (iii) appliquée à la représentation (5.11) signifie, en effet (cf. section 3), que $\nabla^d Y \underset{(h)}{\not\rightarrow} \nabla^d X$. En ce sens, la caractérisation (iii) donne une relation d'équivalence entre non-causalité pour des variables de « stock » et non-causalité pour les flux correspondants (variables différenciées). Il convient, cependant, de rappeler que cette équivalence n'est obtenue que pour *un choix particulier des conditions initiales* et pour un modèle de type ARIMA(p, d, q) (où les relations de coïntégration sont exclues).

ANNEXE

– **Preuve de la Proposition 2 :** Si $Y \underset{h}{\not\longrightarrow} X$, on a par définition l'identité

$$P[X(t+h) \mid I(t) + Y(\omega, t]] = P[X(t+h) \mid I(t)] \, , \forall t > \omega \, ;$$

et donc chaque composante de $P[X(t+h) \mid I(t) + Y(\omega, t]]$ est un élément de $I(t) \subseteq I_{)j(}(t)$ qui coïncide avec sa propre projection sur $I_{)j(}(t)$:

$$P[X(t+h) \mid I(t) + Y(\omega, t]] = P\Big[P[X(t+h) \mid I(t) + Y(\omega, t]] \;\Big|\; I_{)j(}(t)\Big] \, , \forall t > \omega \, .$$

L'identité (2.4) est alors une simple conséquence de la propriété bien connue des « *projections itérées* » : $I_{)j(}(t) \subseteq I(t) + Y(\omega, t]$ implique

$$P\Big[P[X(t+h) \mid I(t) + Y(\omega, t]] \;\Big|\; I_{)j(}(t)\Big] = P[X(t+h) \mid I_{)j(}(t)] \, .$$

En revanche, on peut avoir (2.4) pour $j = 1, 2, \dots, p$ sans avoir (2.3), puisque (2.4) est, par exemple, assuré dès que les p composantes de Y ($p > 1$) sont identiques, alors que ceci n'entraîne évidemment pas que $Y \underset{h}{\not\longrightarrow} X$ relativement à I. Q.E.D.

– **Preuve de la Proposition 3 :** Pour vérifier que (i) implique (ii), il suffit, en effet, d'appliquer le même argument de projections itérées que pour la preuve de la proposition 2, en remplaçant $I_{)j(}(t)$ par $I(t) + y_j(\omega, t]$. Mais, réciproquement, si on a (ii), c'est-à-dire

$$P[X(t+h) \mid I(t) + y_j(\omega, t]] = P[X(t+h) \mid I(t)]$$

pour tout $j = 1, 2, \dots, p$, on peut affirmer que les composantes de $X(t+h) - P[X(t+h) \mid I(t)]$ sont orthogonales à chacun des espaces $I(t) + y_j(\omega, t]$, $j = 1, 2, \dots, p$ et donc aussi à l'espace de Hilbert $I(t) + Y(\omega, t]$ engendré par ces sous-espaces. Ceci prouve bien que $P[X(t+h) \mid I(t)] = P[X(t+h) \mid I(t) + Y(\omega, t]]$, c'est-à-dire $Y \underset{h}{\not\longrightarrow} X$ relativement à I. Finalement, l'équivalence entre (ii) et (iii) est une conséquence directe de la proposition 1. Q.E.D.

– **Preuve de la proposition 4 :** Il s'agit, en fait, de démontrer que si $Y \underset{1}{\not\longrightarrow} X$, alors, pour tout entier positif h, $Y \underset{(h)}{\not\longrightarrow} X$. La preuve se fait par récurrence sur h. On suppose que $Y \underset{(h)}{\not\longrightarrow} X$, et il s'agit de montrer que $Y \underset{h+1}{\not\longrightarrow} X$. D'après le principe des projections itérées,

$$P[X(t+h+1) \mid I(t) + Y(\omega, t]] = P\Big[P[X(t+h+1) \mid I(t+h) + Y(\omega, t+h]] \;\Big|\; I(t) + Y(\omega, t]\Big] \, ,$$

$\forall t > \omega$. Or, puisque $Y \underset{1}{\not\longrightarrow} X$,

$$P[X(t+h+1) \mid I(t+h) + Y(\omega, t+h]] = P[X(t+h+1) \mid I(t+h)] \, , \forall t > \omega \, .$$

Mais $P[X(t+h+1) \mid I(t+h)]$ est un vecteur $m \times 1$ d'éléments de

$$I(t+h) = E + X(\omega, t+h] = E + X(\omega, t] + X[t+1, t+h] = I(t) + X[t+1, t+h]$$

et donc se décompose en une somme $a(t) + b(t)$, où les composantes du vecteur $a(t)$ appartiennent à $I(t)$ et celles de $b(t)$ à $X[t+1, t+h]$. Donc

$$\begin{aligned} P\Big[P[X(t+h+1) \mid I(t+h) + Y(\omega, t+h]] \;\Big|\; I(t) + Y(\omega, t]\Big] &= P[a(t) \mid I(t) + Y(\omega, t]] + P[b(t) \mid I(t) + Y(\omega, t]] \\ &= a(t) + P[b(t) \mid I(t)] \, , \forall t > \omega \, , \end{aligned}$$

puisque $I(t) \subseteq I(t) + Y(\omega, t]$ et $X[t+1, t+h]$ est l'espace de Hilbert engendré par les composantes $x_i(\tau)$, $i = 1, 2, \dots, m$, $t+1 \leq \tau \leq t+h$, dont chacune vérifie (par l'hypothèse de récurrence $Y \underset{(h)}{\not\longrightarrow} X$)

$$P[x_i(\tau) \mid I(t) + Y(\omega, t]] = P[x_i(\tau) \mid I(t)] \, , \forall t > \omega \, .$$

Par conséquent, chaque composante de $P[X(t+h+1) \mid I(t) + Y(\omega, t]]$ appartient à $I(t)$, et puisque $I(t) \subseteq I(t) + Y(\omega, t]$,

$$P[X(t+h+1) \mid I(t) + Y(\omega, t]] = P[X(t+h+1) \mid I(t)] \, , \forall t > \omega \, ,$$

ce qui prouve bien que $Y \underset{(h+1)}{\not\longrightarrow} X$. On a donc (i) $\Leftrightarrow$ (ii). L'équivalence entre (ii) et (iii) découle trivialement de la définition 2. Q.E.D.

– Preuve de la proposition 6 : Soit $t \in \mathbb{Z}$ et $t > \omega$. La preuve se fait par récurrence sur h. D'après (3.2),

$$W(t+1) = a(t+1) + \sum_{j=1}^{\infty} \pi_j W(t+1-j),$$

ce qui donne bien (3.6) pour $h = 1$ avec $\pi_j^{(1)} = \pi_j$, car $a(t+1)$ est orthogonal à l'espace engendré par $\{W(\tau) : \tau \le t\}$ et donc $P[a(t+1) \mid E + W(-\infty, t]] = 0$. Supposons que les égalités (3.6) et (3.7) soient vérifiées jusqu'à h, où $h \ge 1$. Alors, d'après (3.2),

$$W(t+h+1) = a(t+h+1) + \sum_{j=1}^{\infty} \pi_j W(t+h+1-j),$$

et donc, en notant $\bar{I}(t) = I(t) + Y(-\infty, t)$,

$$P[W(t+h+1) \mid \bar{I}(t)] = \sum_{j=1}^{\infty} \pi_j P[W(t+h+1-j) \mid \bar{I}(t)]$$

$$= \sum_{j=1}^{h} \pi_j P[W(t+h+1-j) \mid \bar{I}(t)] + \sum_{j=h+1}^{\infty} \pi_j W(t+h+1-j).$$

Mais, d'après l'hypothèse de récurrence, on a pour $j = 1, 2, \dots, h$,

$$P[W(t+h+1-j) \mid \bar{I}(t)] = \sum_{k=1}^{\infty} \pi_k^{(h+1-j)} W(t+1-k),$$

d'où

$$P[W(t+h+1) \mid \bar{I}(t)] = \sum_{\ell=1}^{h} \pi_{h+1-\ell} \sum_{k=1}^{\infty} \pi_k^{(\ell)} W(t+1-k) + \sum_{k=1}^{\infty} \pi_{k+h} W(t+1-k).$$

On obtient donc bien que

$$\text{(A.1)} \quad P[W(t+h+1) \mid \bar{I}(t)] = \sum_{j=1}^{\infty} \pi_j^{(h+1)} W[t+1-j],$$

avec

$$\pi_j^{(h+1)} = \pi_{j+h} + \sum_{\ell=1}^{h} \pi_{h-\ell+1} \pi_j^{(\ell)},$$

ce qui démontre (3.6) et (3.7) par récurrence. Comme (3.7) est valide pour tout processus de la forme (3.2) sans égard à la matrice de covariance de $a(t)$, considérons un processus $W(t)$ qui satisfait (3.2) pour $t \ge 1$ avec $V[a(t)] = \Sigma$ tel que $\det(\Sigma) \ne 0$ et $W(t) \equiv 0$ pour $t \le 0$. Il est clair qu'un tel processus existe toujours. Alors, d'après le principe des projections itérées, on a pour tout $t \in \mathbb{N}$ et $h \in \mathbb{N}$,

$$\text{(A.2)} \quad P[W(t+h+1) \mid \bar{I}(t)] = P\Big[P[W(t+h+1) \mid \bar{I}(t+1)] \mid \bar{I}(t)\Big]$$

$$= P\left[\sum_{j=1}^{\infty} \pi_j^{(h)} W(t+2-j) \mid \bar{I}(t)\right]$$

$$= \pi_1^{(h)} P[W(t+1) \mid \bar{I}(t)] + \sum_{k=1}^{\infty} \pi_{k+1}^{(h)} W(t+1-k)$$

$$= \sum_{j=1}^{\infty} \left[\pi_1^{(h)} \pi_j + \pi_{j+1}^{(h)}\right] W(t+1-j).$$

En soustrayant (A.2) de (A.1), on voit que

$$\text{(A.3)} \quad \sum_{j=1}^{\infty} \left[\pi_j^{(h+1)} - \pi_1^{(h)} \pi_j - \pi_{j+1}^{(h)}\right] W(t+1-j) = 0, \ \forall t, h \in \mathbb{N}.$$

En multipliant (A.3) par $a(t)$ et en prenant l'espérance mathématique pour $t > 1$, on trouve

$$[\pi_1^{(h+1)} - \pi_1^{(h)} \pi_1 - \pi_2^{(h)}] \Sigma = 0, \ \forall h \in \mathbb{N}$$

d'où, comme la matrice Σ est inversible,

$$\pi_1^{(h+1)} - \pi_1^{(h)} \pi_1 - \pi_2^{(h)} = 0, \ \forall h \in \mathbb{N}.$$

On a donc, en fait,

$$\sum_{j=2}^{\infty} [\pi_1^{(h+1)} - \pi_1^{(h)} \pi_j - \pi_{j+1}^{(h)}] W(t+1-j) = 0, \ \forall t, h \in \mathbb{N}.$$

Multipliant par $a(t-1)$, pour $t > 2$, et prenant l'espérance mathématique, on trouve de même

$$\pi_2^{(h+1)} - \pi_1^{(h)}\pi_2 - \pi_3^{(h)} = 0 \,, \forall h \in \mathbb{N}$$

et, de proche en proche,

$$\pi_j^{(h+1)} - \pi_1^{(h)}\pi_j - \pi_{j+1}^{(h)} = 0 \,, \forall j \,, h \in \mathbb{N} \,.$$

Ceci achève la démonstration de (3.8) par récurrence. Q.E.D.

- Preuve de la proposition 7 : Soit $t \in \mathbb{Z}$, $t > \omega$ et notons $I(t) = I_{XZ}(t)$. On déduit de la proposition 6 que

$$P[X(t+h) \mid I(t) + Y(-\infty, t]] = \sum_{j=1}^{\infty} \left[\pi_{XXj}^{(h)} X(t+1-j) + \pi_{XYj}^{(h)} Y(t+1-j) + \pi_{XZj}^{(h)} Z(t+1-j)\right].$$

La condition (3.10) implique donc que

$$P[X(t+h) \mid I(t) + Y(-\infty, t]] = \sum_{j=1}^{\infty} \left[\pi_{XXj}^{(h)} X(t+1-j) + \pi_{XZj}^{(h)} Z(t+1-j)\right]$$

et $P[X(t+h) \mid I(t) + Y(-\infty, t]]$ a par conséquent toutes ses composantes dans l'espace de Hilbert $I(t) = E + X(-\infty, t] + Z(-\infty, t]$. C'est dire que

$$P[X(t+h) \mid I(t) + Y(-\infty, t]] = P[X(t+h) \mid I(t)] \,.$$

Ainsi, la condition (3.10) est suffisante pour que $Y \underset{h}{\not\longrightarrow} X$ relativement à I.

Supposons maintenant que les matrices $E[a(t)\,a(t)']$ sont non singulières pour $t > \omega$. Si $Y \underset{h}{\not\longrightarrow} X$ relativement à I, toutes les composantes de $P[X(t+h) \mid I(t) + Y(-\infty, t]]$ appartiennent à l'espace de Hilbert $I(t)$, pour $t > \omega$. C'est dire que $P[X(t+h) \mid I(t) + Y(-\infty, t]]$ qui s'écrit

$$P[X(t+h) \mid I(t) + Y(-\infty, t]] = \sum_{j=1}^{\infty} \pi_{X\cdot j}^{(h)} W(t+1-j)$$

où $\pi_{X\cdot j}^{(h)} = \left[\pi_{XXj}^{(h)}, \pi_{XYj}^{(h)}, \pi_{XZj}^{(h)}\right]$, s'écrit aussi comme limite en moyenne quadratique d'une suite

$$U_T = \sum_{j=1}^{T} \phi_j^{(T)} W(t+1-j) \,, T \in \mathbb{N} \,,$$

où chaque U_T a toutes ses composantes dans $I(t)$:

$$U_T = \sum_{j=1}^{T} \left[\phi_{Xj}^{(T)} X(t+1-j) + \phi_{Zj}^{(T)} Z(t+1-j)\right].$$

Par conséquent, en définissant $\phi_j^{(T)} = 0$, pour $j > T$, et

$$\overline{U}_T(t) \equiv P[X(t+h) \mid I(t) + Y(-\infty, t]] - U_T = \sum_{j=1}^{\infty} \left[\pi_{X\cdot j}^{(h)} - \phi_j^{(T)}\right] W(t+1-j) \,,$$

le vecteur $\overline{U}_T(t)$ converge en moyenne quadratique vers zéro et donc

$$E[\overline{U}_T(t)\,a(t)'] = \left[\pi_{X\cdot 1}^{(h)} - \phi_1^{(T)}\right] E[a(t)\,a(t)'] \underset{T\to\infty}{\longrightarrow} 0 \,,$$

car $E[W(t)\,a(t)'] = E[a(t)\,a(t)']$ et $E[W(t+1-j)\,a(t)'] = 0$ pour $j \geq 2$. Comme la matrice $E[a(t)\,a(t)']$ est non singulière, on en déduit que $\pi_{X\cdot j}^{(h)} - \phi_j^{(T)} \underset{T\to\infty}{\longrightarrow} 0$. Comme $\phi_1^{(T)} = \left[\phi_{X1}^{(T)}, 0, \phi_{Z1}^{(T)}\right]$, ceci implique que $\pi_{XY1}^{(h)} = 0$. On voit donc que $\sum_{j=2}^{\infty} \left[\pi_{X\cdot j}^{(h)} - \phi_j^{(T)}\right] W(t+1-j)$ converge en moyenne quadratique vers zéro et un raisonnement analogue (avec $t-1 > \omega$) permet de conclure que $\pi_{XY2}^{(h)} = 0$. De proche en proche, on montre ainsi que $\pi_{XYj}^{(h)} = 0$ pour tout $j = 1, 2, 3, \ldots$. Q.E.D.

- Preuve du théorème 1 : L'équivalence entre (i) et (ii) résulte immédiatement des propositions 7 et 8. On démontre l'équivalence entre (i) et (iii) par récurrence sur h. En effet, l'équivalence est assurée pour $h = 2$, puisqu'alors les critères (ii) et (iii) sont identiques. On suppose donc cette équivalence assurée jusqu'à l'ordre h (l'hypothèse de récurrence) et on veut montrer que $Y \underset{(h+1)}{\not\longrightarrow} X$ si et seulement si

$$\pi_{XYj} = 0 \,, \forall j \geq 1 \,, \text{ et } R_{XZ}^{(k)} \pi_{ZYj} = 0 \,, k = 1, 2, \ldots, h \,, \forall j \geq 1 \,.$$

Comme $Y \underset{(h+1)}{\not\longrightarrow} X$ si et seulement si $Y \underset{(h)}{\not\longrightarrow} X$ et $Y \underset{h+1}{\not\longrightarrow} X$, il s'agit, en fait, de vérifier que, *sachant que* $Y \underset{(h)}{\not\longrightarrow} X$, $Y \underset{h+1}{\not\longrightarrow} X$ si et seulement si

$$R_{XZ}^{(h)} \pi_{ZYj} = 0 \, , \forall j \geq 1 \, .$$

Compte tenu du critère (ii), il suffit donc de montrer que si $Y \underset{(h)}{\not\longrightarrow} X$, alors

$$\pi_{XZ1}^{(h)} \pi_{XYj} = R_{XZ}^{(h)} \pi_{ZYj} \, , \forall j \geq 1 \, .$$

On va, en fait, démontrer le lemme suivant.

Lemme : Sous les hypothèses du théorème 1, la proposition suivante (P_h) est vraie pour tout entier h supérieur à 1 : si $Y \underset{(h)}{\not\longrightarrow} X$ et si p est un entier tel que $2 \leq p \leq h$, alors

$$\pi_{XZ1}^{(h)} \pi_{ZYj} = \sum_{\ell=1}^{p} \pi_{XZ\ell}^{(h-p+1)} \left\{ \sum_{J(p-\ell)} \left[\prod_{i=1}^{n_{p-\ell}} \pi_{ZZi}^{n_i} \right] \right\} \pi_{ZYj}$$

pour tout entier positif j.

La proposition (P_h), appliquée avec p = h, donne bien, en particulier, l'égalité cherchée. La démonstration du lemme se fait en deux étapes.

1re étape : on vérifie le résultat pour p = 2 (pour tout $h \geq 2$). Il s'agit de montrer que si $Y \underset{(h)}{\not\longrightarrow} X$,

$$\pi_{XZ1}^{(h)} \pi_{ZYj} = \left[\pi_{XZ1}^{(h-1)} \pi_{ZZ1} + \pi_{XZ2}^{(h-1)} \right] \pi_{ZYj} \, .$$

Or, d'après la formule de récurrence (3.8),

$$\pi_1^{(h)} = \pi_1^{(h-1)} \pi_1 + \pi_2^{(h-1)}$$

et en particulier,

$$\pi_{XZ1}^{(h)} = \pi_{XX1}^{(h-1)} \pi_{XZ1} + \pi_{XY1}^{(h-1)} \pi_{YZ1} + \pi_{XZ1}^{(h-1)} \pi_{ZZ1} + \pi_{XZ2}^{(h-1)} \, .$$

Comme $Y \underset{h-1}{\not\longrightarrow} X$, $\pi_{XY1}^{(h-1)} = 0$ et donc

$$\pi_{XZ1}^{(h)} \pi_{ZYj} = \left[\pi_{XX1}^{(h-1)} \pi_{XZ1} + \pi_{XZ1}^{(h-1)} \pi_{ZZ1} \right] \pi_{ZYj} + \pi_{XZ2}^{(h-1)} \pi_{ZYj} \, .$$

Mais, comme $Y \underset{2}{\not\longrightarrow} X$ (puisque $h \geq 2$), on sait d'après le point (ii) du théorème 1 que $\pi_{XZ1} \pi_{ZYj} = 0$, d'où le résultat annoncé.

2e étape : on démontre la propriété (P_h) par récurrence sur h. Pour h = 2, la propriété (P_h) résulte de la première étape ($2 \leq p \leq h \Leftrightarrow p = 2$). Supposons que pour $h \geq 3$, les propriétés (P_2), (P_3), ... , (P_{h-1}) soient vraies et démontrons, sous cette hypothèse de récurrence, la propriété (P_h). En supposant $Y \underset{(h)}{\not\longrightarrow} X$, on doit établir l'égalité annoncée pour p = 2, 3, ... , h. Comme on l'a déjà établi pour p = 2 (1re étape), il suffit, en vertu du principe de récurrence (sur p), de l'établir pour p + 1 en la supposant vraie aux ordres 2, 3, ... , p < h. En particulier, à l'ordre p, on sait que

$$\pi_{XZ1}^{(h)} \pi_{ZYj} = \left[\sum_{\ell=1}^{p} \pi_{XZ\ell}^{(h-p+1)} \sum_{J(p-\ell)} \prod_i \pi_{ZZi}^{n_i} \right] \pi_{ZYj} \, ,$$

pour tout entier positif j, où pour alléger la notation nous écrivons

$$\sum_{J(\ell)} \prod_i \pi_{ZZi}^{n_i} = \sum_{J(\ell)} \left[\prod_{i=1}^{\ell} \pi_{ZZi}^{n_i} \right]$$

pour tout entier ℓ. Mais, d'après (3.8),

$$\pi_\ell^{(h-p+1)} = \pi_{\ell+1}^{(h-p)} + \pi_1^{(h-p)} \pi_\ell$$

et, en particulier,

$$\pi_{XZ\ell}^{(h-p+1)} = \pi_{XX1}^{(h-p)} \pi_{XZ\ell} + \pi_{XY1}^{(h-p)} \pi_{YZ\ell} + \pi_{XZ1}^{(h-p)} \pi_{ZZ\ell} + \pi_{XZ,\ell+1}^{(h-p)} \, .$$

Comme $Y \underset{(h)}{\not\longrightarrow} X$, $\pi_{XY1}^{(h-p)} = 0$ et donc

$$\left[\sum_{\ell=1}^{p} \pi_{XZ\ell}^{(h-p+1)} \sum_{J(p-\ell)} \prod_i \pi_{ZZi}^{n_i} \right] \pi_{ZYj} = \left\{ \sum_{\ell=1}^{p} \left[\pi_{XZ,\ell+1}^{(h-p)} + \pi_{XZ1}^{(h-p)} \pi_{ZZ\ell} \right] \sum_{J(p-\ell)} \prod_i \pi_{ZZi}^{n_i} \right\} \pi_{ZYj}$$

$$+ \pi_{XX1}^{(h-p)} \left[\sum_{\ell=1}^{p} \pi_{XZ\ell} \left[\sum_{J(p-\ell)} \prod_i \pi_{ZZi}^{n_i} \right] \right] \pi_{ZYj} \, .$$

On va montrer que le second terme de cette somme de deux termes est nul en utilisant la propriété (P_p) qui est vraie par hypothèse de récurrence (car $p \le h-1$) et qui peut être appliquée (car $Y \underset{(p)}{\not\longrightarrow} X$). La propriété (P_p) nous dit que pour $2 \le q \le p$,

$$\pi_{XZ1}^{(p)} \pi_{ZYj} = \left[\sum_{\ell=1}^{q} \pi_{XZ\ell}^{(p-q+1)} \sum_{J(q-\ell)} \prod_i \pi_{ZZi}^{n_i} \right] \pi_{ZYj} .$$

En particulier pour $q = p$, on obtient

$$\pi_{XZ1}^{(p)} \pi_{ZYj} = \left[\sum_{\ell=1}^{p} \pi_{XZ\ell} \sum_{J(p-\ell)} \prod_i \pi_{ZZi}^{n_i} \right] \pi_{ZYj} .$$

Mais comme $Y \underset{(p+1)}{\not\longrightarrow} X$ (puisque $p+1 \le h$), on sait, d'après le point (ii) du théorème 1, que $\pi_{XZ1}^{(p)} \pi_{ZYj} = 0$, d'où la nullité du terme annoncé. Par conséquent,

$$\pi_{XZ1}^{(h)} \pi_{ZYj} = \left\{ \sum_{\ell=1}^{p} \left[\pi_{XZ,\ell+1}^{(h-p)} + \pi_{XZ1}^{(h-p)} \pi_{ZZ\ell} \right] \sum_{J(p-\ell)} \prod_i \pi_{ZZi}^{n_i} \right\} \pi_{ZYj}$$

$$= \left[\sum_{\ell=2}^{p+1} \pi_{XZ\ell}^{(h-p)} \sum_{J(p+1-\ell)} \prod_i \pi_{ZZi}^{n_i} \right] \pi_{ZYj} + \pi_{XZ1}^{(h-p)} \left[\sum_{\ell=1}^{p} \sum_{J(p-\ell)} \pi_{ZZ\ell} \prod_i \pi_{ZZi}^{n_i} \right] \pi_{ZYj} .$$

Mais il est clair par la définition de $J(p)$ que

$$\sum_{\ell=1}^{p} \sum_{J(p-\ell)} \prod_i \pi_{ZZ\ell} \pi_{ZZi}^{n_i} = \sum_{J(p)} \prod_i \pi_{ZZi}^{n_i} ,$$

d'où

$$\pi_{XZ1}^{(h)} \pi_{ZYj} = \left[\sum_{\ell=1}^{p+1} \pi_{XZ\ell}^{(h-p)} \sum_{J(p+1-\ell)} \prod_i \pi_{ZZi}^{n_i} \right] \pi_{ZYj} ,$$

ce qui démontre bien le résultat énnoncé à l'ordre $p+1$. La propriété (P_h) est donc bien établie par récurrence.
Q.E.D.

- Preuve des propositions 9 et 10 : Même si la proposition 9 a été énoncée avant la proposition 10 à partir d'un argument intuitif, on en donne ici une démonstration rigoureuse à partir du résultat de la proposition 10 démontré au préalable.

a) <u>Preuve de la proposition 10 :</u> Compte tenu de (3.11), on a :

$$\pi_{ZZ}(z) = I_r - \sum_{k=1}^{\infty} \pi_{ZZk} z^k , \ \pi_{XZ}(z) = -\sum_{j=1}^{\infty} \pi_{XZj} z^j$$

et donc

$$[\pi_{ZZ}(z)]^{-1} = I_r + \sum_{\ell=1}^{\infty} \left[\sum_{k=1}^{\infty} \pi_{ZZk} z^k \right]^{\ell} = \sum_{k=0}^{\infty} \left\{ \sum_{J(k)} \left[\prod_{i=1}^{k} \pi_{ZZi}^{n_i} \right] \right\} z^k$$

où $J(k)$, $k \in \mathbb{N}_0$, est défini conformément aux notations du théorème 1,

$$-\pi_{XZ}(z) [\pi_{ZZ}(z)]^{-1} = \sum_{k=1}^{\infty} \left\{ \sum_{\ell=1}^{k} \pi_{XZ\ell} \left[\sum_{J(k-\ell)} \left[\prod_{i=1}^{k-\ell} \pi_{ZZi}^{n_i} \right] \right] \right\} z^k = \sum_{k=1}^{\infty} R_{XZ}^{(k)} z^k = R_{XZ}(z) .$$

Q.E.D.

b) <u>Preuve de la proposition 9 :</u> D'après le théorème 1, une condition nécessaire et suffisante pour que Y ne cause pas X à tout horizon est que

$$\pi_{XYj} = 0 , \ \forall j \ge 1 , \text{ et } R_{XZ}^{(k)} \pi_{ZYj} = 0, \ \forall j \ge 1 , \ \forall k \ge 1 .$$

Ceci s'écrit aussi en termes de fonctions génératrices :

$$\pi_{XY}(z) \equiv 0 \quad \text{et} \quad R_{XZ}(z) \pi_{ZYj} \equiv 0, \ \forall j \ge 1 .$$

En multipliant $R_{XZ}(z) \pi_{ZYj}$ par z^j et en sommant sur $j \in \mathbb{N}$, on en déduit qu'une condition nécessaire pour que Y ne cause pas X à tout horizon est que

$$\pi_{XY}(z) \equiv 0 \quad \text{et} \quad R_{XZ}(z) \pi_{ZY}(z) \equiv 0$$

ou encore, compte tenu de la proposition 10,

$$\pi_{XY}(z) \equiv 0 \text{ et } \pi_{XZ}(z) \pi_{ZZ}^{-1}(z) \pi_{XY}(z) \equiv 0 .$$

D'où la condition (i) de la proposition 9.

Pour montrer l'équivalence entre (i) et (ii), il suffit d'utiliser les formules usuelles pour l'inversion de matrices partitionnées pour obtenir la formule (en omettant le symbole z pour alléger la notation) :

$$\psi_{XY} = -\pi_{XX\cdot Z}^{-1}[\pi_{XY} - \pi_{XZ}\,\pi_{ZZ}^{-1}\,\pi_{ZY}]\,\tilde{\pi}_{YY}^{-1}$$

où $\pi_{XX\cdot Z} = \pi_{XX} - \pi_{XZ}\,\pi_{ZZ}^{-1}\,\pi_{ZX}$,

$$\tilde{\pi}_{YY} = \pi_{YY} - [\pi_{YX}, \pi_{YZ}]\,A\begin{bmatrix}\pi_{XY}\\ \pi_{ZY}\end{bmatrix},\ A = \begin{bmatrix}\pi_{XX} & \pi_{XZ}\\ \pi_{ZX} & \pi_{ZZ}\end{bmatrix},$$

et où les inverses des matrices π_{ZZ}, A, $\pi_{XX\cdot Z}$ et $\tilde{\pi}_{YY}$ existent toutes dans un disque suffisamment petit centré à l'origine (disons pour $|z| < \delta$, où $0 < \delta < 1$); voir Dufour et Tessier (1991, Proposition 1) pour un développement analogue. Pour $|z| < \delta$, on voit alors que

$$\pi_{XY}(z) \equiv 0 \ \text{ et } \ \pi_{XZ}(z)\,\pi_{ZZ}(z)^{-1}\,\pi_{ZY}(z) \equiv 0 \Leftrightarrow \pi_{XY}(z) \equiv 0 \ \text{ et } \ \psi_{XY}(z) \equiv 0,$$

ce qui est bien l'équivalence entre (i) et (ii). L'équivalence entre (ii) et (iii) se déduit de celle entre (i) et (ii) en permutant les rôles de π et ψ, puisque (ii) leur fait jouer un rôle symétrique.

Q.E.D.

- Preuve des propositions 12 et 13 :

On s'appuiera, pour ces démonstrations, sur un lemme d'algèbre que le lecteur vérifiera aisément [pour une démonstration, voir Dufour et Tessier (1992)]. Ce lemme généralise une propriété utilisée par Lütkepohl (1990)

Lemme : Soit $f(z) = \sum_{j=0}^{\infty} a_j z^j$ une série entière en $z \in \mathbb{C}$ (avec $a_j \in \mathbb{C}$ pour $j \geq 0$) qui est convergente pour $|z| < \delta$, avec $\delta > 0$, et telle que

$$\sum_{j=0}^{\infty} a_j z^j = \left[\sum_{k=0}^{p} b_k z^k\right]\left[\sum_{\ell=0}^{\infty} c_\ell z^\ell\right]$$

pour $|z| < \delta$, où $\sum_{k=0}^{p} b_k z^k$ est un polynôme de degré p et $\sum_{\ell=0}^{\infty} c_\ell z^\ell$, avec $c_0 = 1$, est une autre série entière qui est convergente pour $|z| < \delta$. Alors

$$a_j = 0,\ \forall j \geq 0 \Leftrightarrow a_j = 0,\ \text{pour } j = 0, 1, \dots, p.$$

Pour démontrer la proposition 12, il s'agit, compte tenu du théorème 1, de montrer que : si $R_{XZ}^{(k)}\,\pi_{ZYj} = 0$, $\forall j \geq 1$, $k = 1, 2, \dots, rP$, alors

$$R_{XZ}^{(k)}\,\pi_{ZYj} = 0,\ \forall j \geq 1,\ \forall k \geq 1,$$

la dernière assertion étant équivalente à

$$R_{XZ}(z)\,\pi_{ZYj} \equiv 0,\ \forall j \geq 1.$$

Mais, d'après la proposition 10,

$$R_{XZ}(z)\,\pi_{ZYj} = -\pi_{XZ}(z)\,\pi_{ZZ}(z)^{-1}\,\pi_{ZYj} = -[\det \pi_{ZZ}(z)]^{-1}\,\pi_{XZ}(z)\,\pi_{ZZ}^{*}(z)\,\pi_{ZYj}$$

où $\pi_{ZZ}^{*}(z)$ est la matrice transposée de la matrice de cofacteurs de $\pi_{ZZ}(z)$. $\pi_{ZZ}(z)$ étant une matrice carrée de taille r dont les coefficients sont des polynômes en z de degré inférieur ou égal à P par hypothèse, $\pi_{ZZ}^{*}(z)$ est une matrice dont les coefficients sont des polynômes en z de degré inférieur ou égal à $(r-1)P$.

Considérons alors un couple donné (x_i, y_ℓ) de composantes de $(X', Y')'$, où $1 \leq i \leq n$, $1 \leq \ell \leq P$. On a :

$$R_{x_i Z}(z)\,\pi_{Zy_\ell j} = -[\det \pi_{ZZ}(z)]^{-1}\,\pi_{x_i Z}(z)\pi_{ZZ}^{*}(z)\pi_{Zy_\ell j},$$

où $[\det \pi_{ZZ}(z)]^{-1} = \sum_{\ell=0}^{\infty} c_\ell z^\ell$ est une série formelle avec $c_0 = 1$ et $\pi_{x_i Z}(z)\pi_{ZZ}^{*}(z)\pi_{Zy_\ell j}$ est un polynôme de degré inférieur ou égal à rP. On peut alors affirmer, en vertu du lemme énoncé ci-dessus, que

$$R_{x_i Z}(z)\, \pi_{Zy_\ell j} = 0 \,,\ \forall j \geq 1 \,,$$

si et seulement si, pour tout $j \geq 1$, les monômes de degré plus petit ou égal à rP dans le polynôme $R_{x_i Z}(z)\, \pi_{Zy_\ell j}$ ont des coefficients nuls :

$$R^{(k)}_{x_i Z}\, \pi_{Zy_\ell j} = 0 \,, \text{ pour } k = 1, 2, \ldots, rP \,.$$

Le résultat annoncé résulte alors du théorème 1.

La proposition 13 résulte encore plus immédiatement du lemme énoncé ci-dessus, puisque pour un couple donné (x_i, y_ℓ) de composantes de $(X', Y')'$:

$$\pi_{x_i Z}(z)\, \pi_{ZZ}(z)^{-1}\, \pi_{Zy_\ell}(z) \equiv 0 \Leftrightarrow [\det \pi_{ZZ}(z)]^{-1}\, \pi_{x_i Z}(z)\, \pi^*_{ZZ}(z)\, \pi_{Zy_\ell}(z) \equiv 0$$

ce qui est assuré, en vertu du lemme, dès que les rP+P monômes de plus bas degré sont nuls, car $\pi_{x_i Z}(z)\, \pi^*_{ZZ}(z)\, \pi_{Zy_\ell}(z)$ est un polynôme de degré au plus égal à rP+P. Q.E.D.

- **Preuve de la proposition 14** : On suppose que $Y \underset{(h)}{\longrightarrow} X$ et on considère un entier j tel que $2 \leq j \leq h - 1$. Pour $k = j - 1$, on peut appliquer la formule de récurrence (3.7) :

$$\pi_1^{(k+1)} = \pi_{k+1} + \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, \pi_1^{(\ell)} \,.$$

Autrement dit,

$$\pi_1^{(j)} = \pi_j + \sum_{\ell=1}^{j-1} \pi_{j-\ell}\, \pi_1^{(\ell)} \,.$$

En particulier,

$$\pi^{(j)}_{XZ1} = \pi_{XZj} + \sum_{\ell=1}^{j-1} \left[\pi_{XX,j-\ell}\, \pi^{(\ell)}_{XZ1} + \pi_{XZ,j-\ell}\, \pi^{(\ell)}_{ZZ1}\right]$$

puisque $\pi_{XY,j-\ell} = 0$ (car $Y \underset{1}{\longrightarrow} X$). En multipliant à droite par Q_2', on en déduit que

$$\tilde{\pi}^{(j)}_{XZ_2 1} = \tilde{\pi}_{XZ_2 j} + \sum_{\ell=1}^{j-1} \left[\pi_{XX,j-\ell}\, \tilde{\pi}^{(\ell)}_{XZ_2 1} + \pi_{XZ,j-\ell}\, \pi^{(\ell)}_{ZZ1}\, Q_2'\right] .$$

Mais, puisque $Y \underset{(h)}{\longrightarrow} X$ et $j \leq h - 1$, on sait, d'après le corollaire 1.1, que

$$\tilde{\pi}^{(\ell)}_{XZ_2 1} = 0 \,, \text{ pour } \ell = 1, 2, \ldots, j \,.$$

On a donc

$$0 = \tilde{\pi}_{XZ_2 j} + \sum_{\ell=1}^{j-1} \pi_{XZ,j-\ell}\, \pi^{(\ell)}_{ZZ1}\, Q_2' = \tilde{\pi}_{XZ_2 j} + \sum_{\ell=1}^{j-1} \pi_{XZ,j-\ell}\, [Q_1', Q_2'] \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \pi^{(\ell)}_{ZZ1}\, Q_2'$$

$$= \tilde{\pi}_{XZ_2 j} + \sum_{\ell=1}^{j-1} \left[\tilde{\pi}_{XZ_1,j-\ell}\, \tilde{\pi}^{(\ell)}_{Z_1 Z_2 1} + \tilde{\pi}_{XZ_2,j-\ell}\, \tilde{\pi}^{(\ell)}_{Z_2 Z_2 1}\right] ,$$

d'où le résultat annoncé. Q.E.D.

- **Preuve de la proposition 15** : On suppose que $Y \underset{(h)}{\longrightarrow} X$ et que

$$\sum_{\ell=1}^{j-1} \tilde{\pi}_{XZ_1,j-\ell}\, \tilde{\pi}^{(\ell)}_{Z_1 Z_2 1} = 0 \,, \text{ pour } j = 2, 3, \ldots, h - 1 \,.$$

On a alors, d'après la proposition 14,

$$\tilde{\pi}_{XZ_2 j} = -\sum_{\ell=1}^{j-1} \tilde{\pi}_{XZ_2,j-\ell}\, \tilde{\pi}^{(\ell)}_{Z_2 Z_2 1} \,, \text{ pour } j = 2, 3, \ldots, h - 1 \,.$$

Sachant que $\tilde{\pi}_{XZ_2 1} = 0$ (voir corollaire 1.1), on en déduit, de proche en proche, que

$$\tilde{\pi}_{XZ_2 j} = 0 \,, \text{ pour } j = 2, 3, \ldots, h - 1 \,.$$ Q.E.D.

- **Preuve de la proposition 17** : (i) On suppose que W(t) vérifie (5.1), pour tout $t \in \mathbb{N}_0$, et on montre, par récurrence sur t, que W(t) vérifie (5.2). Pour t = 0,

$$W(0) = a(0) + C(0)\, \overline{W}(-1) = a(0) + B(0)\, \overline{W}(-1) \,.$$

Supposons que la propriété (5.2) soit vraie jusqu'à t. Alors

$$W(t+1) = a(t+1) + \sum_{j=1}^{t+1} \pi_j W(t+1-j) + C(t+1)\,\overline{W}(-1)$$

$$= a(t+1) + \sum_{j=1}^{t+1} \pi_j\, a(t+1-j) + \sum_{j=1}^{t+1} \pi_j \sum_{i=1}^{t+1-j} H_i\, a(t+1-j-i)$$

$$+ \sum_{j=1}^{t+1} \pi_j B(t+1-j)\,\overline{W}(-1) + C(t+1)\,\overline{W}(-1)$$

$$= a(t+1) + \sum_{j=1}^{t+1} \left[\pi_j + \sum_{i=1}^{j-1} \pi_{j-i} H_i\right] a(t+1-j)$$

$$+ \left[C(t+1) + \sum_{j=1}^{t+1} \pi_j B(t+1-j)\right] \overline{W}(-1)\,,$$

ce qui montre bien que

$$W(t+1) = a(t+1) + \sum_{j=1}^{t+1} H_j\, a(t+1-j) + B(t+1)\,\overline{W}(-1)\,,$$

où H_j et $B(t+1)$ sont conformes à (5.3). La propriété (5.2) est donc établie par récurrence.

(ii) On suppose réciproquement que W(t) vérifie (5.2) pour tout $t \in \mathbb{N}_0$, et on montre par récurrence sur t que W(t) vérifie (5.1). Il est clair que (5.1) tient pour t = 0. Supposons que la propriété (5.1) soit vraie jusqu'à t. Alors

$$W(t+1) = a(t+1) + \sum_{j=1}^{t+1} H_j\, a(t+1-j) + B(t+1)\,\overline{W}(-1)$$

$$= a(t+1) + \sum_{j=1}^{t+1} H_j W(t+1-j) - \sum_{j=1}^{t+1} H_j \sum_{k=1}^{t+1-j} \pi_k W(t+1-j-k)$$

$$- \sum_{j=1}^{t+1} H_j\, C(t+1-j)\,\overline{W}(-1) + B(t+1)\,\overline{W}(-1)$$

$$= a(t+1) + \sum_{j=1}^{t+1} \left[H_j - \sum_{k=1}^{j-1} H_{j-k}\, \pi_k\right] W(t+1-j)$$

$$+ \left[B(t+1) - \sum_{j=1}^{t+1} H_j\, C(t+1-j)\right] \overline{W}(-1)\,,$$

ce qui montre bien que

$$W(t+1) = a(t+1) + \sum_{j=1}^{t+1} \pi_j W(t+1-j) + C(t+1)\,\overline{W}(-1)$$

où π_j et C(t + 1) sont conformes à (5.4). La propriété (5.1) est donc établie par récurrence.

Q.E.D.

– **Preuve de la proposition 20** : La démonstration se fait par récurrence sur h, de façon très similaire à la preuve de la proposition 6. La proposition 18 donne le résultat pour h = 1. On suppose alors (hypothèse de récurrence) que (5.6) et (5.7) sont vérifiées jusqu'au rang h, où $h \geq 1$. En notant $\overline{I}(t) = I(t) + Y[0, t]$, on part de la représentation « autorégressive »

$$W(t+h+1) = a(t+h+1) + \sum_{j=1}^{t+h+1} \pi_j W(t+h+1-j) + C(t+h+1)\,\overline{W}(-1)$$

dont on écrit la projection sur $\overline{I}(t)$:

$$P[W(t+h+1) \mid \overline{I}(t)] = \sum_{j=1}^{h} \pi_j P[W(t+h+1-j) \mid \overline{I}(t)]$$

$$+ \sum_{j=h+1}^{t+h+1} \pi_j W(t+h+1-j) + C(t+h+1)\,\overline{W}(-1)$$

et on applique l'hypothèse de récurrence qui permet de remplacer $P[W(t+h+1-j) \mid \bar{I}(t)]$, pour $j = 1, 2, \ldots, h$, par

$$\sum_{\ell=1}^{t+1} \pi_{\ell}^{(h+1-j)} W(t+1-\ell) + C^{(h+1-j)}(t+1)\, \overline{W}(-1).$$

Des regroupements analogues à ceux opérés dans la preuve de la proposition 6 permettent alors d'obtenir

$$P[W(t+h+1) \mid \bar{I}(t)] = \sum_{j=1}^{t+1} \left[\pi_{j+h} + \sum_{\ell=1}^{h} \pi_{h-\ell+1}\, \pi_j^{(\ell)}\right] W(t+1-j)$$

$$+ \left[C(t+h+1) + \sum_{\ell=1}^{h} \pi_{h-\ell+1}\, C^{(\ell)}(t+1)\right] \overline{W}(-1)$$

ce qui donne bien (5.6) et (5.7) au rang $(h+1)$.

Pour chaque entier positif j, comme la suite $\pi_j^{(h)}$, $h \geq 1$, est définie par la récurrence (5.7) [définition identique à (3.7)], cette dernière doit aussi vérifier (5.8a) [identique à (3.8)] d'après le résultat d'équivalence donné par la proposition 6. Il reste à vérifier (5.8b). Pour (5.7), nous avons pour $h = 1$,

$$C^{(2)}(t+1) = C(t+2) + \pi_1\, C(t+1)$$

et (5.8b) se trouve vérifié. Supposons que (5.8b) est vérifié jusqu'au rang h, où $h \geq 1$. D'après le principe des projections itérées, on a :

$$P[W(t+h+1) \mid \bar{I}(t)] = P[P[W(t+h+1) \mid \bar{I}(t+1)] \mid \bar{I}(t)]$$

$$= P\left[\sum_{j=1}^{t+2} \pi_j^{(h)} W(t+2-j) + C^{(h)}(t+2)\, \overline{W}(-1) \mid \bar{I}(t)\right]$$

$$= \pi_1^{(h)} P[W(t+1) \mid \bar{I}(t)] + \sum_{j=2}^{t+2} \pi_j^{(h)} W(t+2-j) + C^{(h)}(t+2)\, \overline{W}(-1)$$

$$= \pi_1^{(h)} \sum_{j=1}^{t+1} \pi_j W(t+1-j) + \pi_1^{(h)} C(t+1)\, \overline{W}(-1) + \sum_{j=2}^{t+2} \pi_j^{(h)} W(t+2-j)$$

$$+ C^{(h)}(t+2)\, \overline{W}(-1)$$

$$= \sum_{j=1}^{t+1} [\pi_1^{(h)} \pi_j + \pi_{j+1}^{(h)}]\, W(t+1-j) + [\pi_1^{(h)} C(t+1) + C^{(h)}(t+2)]\, \overline{W}(-1)$$

$$= \sum_{j=1}^{t+1} \pi_j^{(h+1)} W(t+1-j) + [\pi_1^{(h)} C(t+1) + C^{(h)}(t+2)]\, \overline{W}(-1)$$

où la dernière identité provient de (5.8a). Comme, d'après (5.6), on peut aussi écrire

$$P[W(t+h+1) \mid \bar{I}(t)] = \sum_{j=1}^{t+1} \pi_j^{(h+1)} W(t+1-j) + C^{(h+1)}(t+1)\, \overline{W}(-1),$$

on en déduit bien le résultat cherché en soustrayant l'une de l'autre les deux expressions pour $P[W(t+h+1) \mid \bar{I}(t)]$. Q.E.D.

- **Preuve de la proposition 21** : La démonstration est très similaire à celle de la proposition 7. La condition (5.9) implique à l'évidence que $P[X(t+h) \mid I(t) + Y(0,t)]$, calculé par (5.6), a toutes ses composantes dans l'espace de Hilbert $I(t) = X[0,t] + Z[0,t] + \overline{W}_d[-1]$, où on note $\overline{W}_d[-1]$ l'espace de Hilbert engendré par les composantes de $\overline{W}_d(-1)$. La condition (5.9) est donc suffisante pour que $Y \underset{h}{\longrightarrow} X$ relativement à I.

Supposons maintenant que les matrices $E[a(t)\, a(t)']$ sont non singulières pour $t \geq 0$. Si $Y \underset{h}{\longrightarrow} X$ relativement à I, toutes les composantes de $P[X(t+h) \mid I(t) + Y[0,t]]$ appartiennent à l'espace de Hilbert $I(t)$. C'est dire que $P[X(t+h) \mid I(t) + Y[0,t]]$ qui s'écrit, avec des notations évidentes,

$$P[X(t+h) \mid I(t) + Y[0,t]] = \sum_{j=1}^{t+1} \pi_{X \cdot j}^{(h)} W(t+1-j) + C_{X \cdot}^{(h)}(t+1)\, \overline{W}(-1)$$

s'écrit aussi

$$P[X(t+h) \mid I(t) + Y[0,t]] = \sum_{j=1}^{t+1} [\phi_{Xj} X(t+1-j) + \phi_{Zj} Z(t+1-j)] + D_X^{(h)}(t+1)\, \bar{W}(-1)$$

$$= \sum_{j=1}^{t+1} \phi_j W(t+1-j) + D_X^{(h)}(t+1)\, \bar{W}(-1)$$

avec $\phi_j = [\phi_{Xj}, 0, \phi_{Zj}]$. En comparant les deux expressions de $P[X(t+h) \mid I(t) + Y(0,t]$ ainsi obtenues, on en déduit que

$$0 = \sum_{j=1}^{t+1} \left[\pi_{X\cdot j}^{(h)} - \phi_j\right] W(t+1-j) + \left[C_X^{(h)}(t+1) - D_X^{(h)}(t+1)\right] \bar{W}(-1) .$$

En multipliant par la droite des deux côtés de la dernière identité par $a(t)'$ et en prenant l'espérance mathématique, on voit alors que

$$\left[\pi_{X\cdot 1}^{(h)} - \phi_1\right] E[a(t)\, a(t)'] = 0$$

car $a(t)$ est non corrélé avec $a(t+1-j)$, $j = 2, 3, \ldots, t+1$, et $\bar{W}(-1)$. Comme la matrice $E[a(t)\, a(t)']$ est non singulière, on en déduit que $\pi_{X\cdot 1}^{(h)} - \phi_1 = 0$. Comme $\phi_1 = [\phi_{X1}, 0, \phi_{Z1}]$, ceci implique que $\pi_{XY1}^{(h)} = 0$. On a donc aussi

$$0 = \sum_{j=2}^{t+1} \left[\pi_{X\cdot j}^{(h)} - \phi_j\right] W(t+1-j) + \left[C_X^{(h)}(t+1) - D_X^{(h)}(t+1)\right] \bar{W}(-1)$$

et un raisonnement analogue permet de conclure que $\pi_{XY2}^{(h)} = 0$ (en prenant $t \geq j$). De proche en proche, on montre ainsi que $\pi_{XYj}^{(h)} = 0$ pour tout $j = 1, 2, \ldots$. Q.E.D.

- **Preuve de la proposition 22** : La preuve se fait par récurrence sur h. Pour $h = 0$, la propriété annoncée correspond simplement à la représentation (5.11), compte tenu de la convention $C_d(t, 0)\, \bar{W}_d(-1) = 0$. Supposons maintenant (hypothèse de récurrence) que la propriété soit vraie à l'ordre h, où $0 \leq h < d$. Alors, on peut écrire

$$\nabla^{d-h} \bar{W}(t) = \nabla^{d-h-1} \bar{W}(t) - \nabla^{d-h-1} \bar{W}(t-1)$$

si bien que l'hypothèse de récurrence donne

$$\nabla^{d-h-1}\bar{W}(t) = \nabla^{d-h-1}\bar{W}(t-1) + a(t) + \sum_{j=1}^{t} (\nabla^h \pi_j^*)\,(\nabla^{d-h-1}\bar{W}(t-j) - \nabla^{d-h-1}\bar{W}(t-j-1))$$

$$+ C_d(t,h)\, \bar{W}_d(-1) + \sum_{j=t+1}^{\infty} \pi_j^*\, \nabla^d \bar{W}(t-j) .$$

Mais

$$\nabla^{d-h-1}\bar{W}(t-1) + \sum_{j=1}^{t} (\nabla^h \pi_j^*)\,(\nabla^{d-h-1}\bar{W}(t-j) - \nabla^{d-h-1}\bar{W}(t-j-1))$$

$$= [I_n + \nabla^h \pi_1^*]\, \nabla^{d-h-1}\bar{W}(t-1) + \sum_{j=2}^{t} (\nabla^h \pi_j^*)\,(\nabla^{d-h-1}\bar{W}(t-j)$$

$$- \sum_{j=2}^{t} (\nabla^h \pi_{j-1}^*)\, \nabla^{d-h-1}\bar{W}(t-j) - (\nabla^h \pi_t^*)\, \nabla^{d-h-1}\bar{W}(-1)$$

$$= \sum_{j=1}^{t} (\nabla^{h+1} \pi_j^*)\,(\nabla^{d-h-1}\bar{W}(t-j) - (\nabla^h \pi_t^*)\, \nabla^{d-h-1}\bar{W}(-1)$$

compte tenu de la convention $I_n = -\nabla^h \pi_0^*$, et

$$C_d(t, h+1)\, \bar{W}_d(-1) = C_d(t,h)\, \bar{W}_d(-1) - (\nabla^h \pi_t^*)\, \nabla^{d-h-1}\bar{W}(-1) .$$

Ainsi, en reportant dans l'expression de $\nabla^{d-h-1}\bar{W}(t)$, on obtient

$$\nabla^{d-h-1}\bar{W}(t) = a(t) + \sum_{j=1}^{t} (\nabla^{h+1}\pi_j^*)\, \nabla^{d-h-1}\bar{W}(t-j) + C_d(t,h)\, \bar{W}_d(-1) - (\nabla^h \pi_t^*)\, \nabla^{d-h-1}\bar{W}(-1)$$

$$+ \sum_{j=t+1}^{\infty} \pi_j^*\, \nabla^d \bar{W}(t-j)$$

$$= a(t) + \sum_{j=1}^{t} (\nabla^{h+1}\pi_j^*)\, \nabla^{d-h-1}\bar{W}(t-j) - C_d(t, h+1)\, \bar{W}_d(-1)$$

$$+ \sum_{j=t+1}^{\infty} \pi_j^*\, \nabla^d \bar{W}(t-j) .$$

La proposition 22 est donc démontrée par récurrence sur h. Q.E.D.

- Preuve de la proposition 23 : Notons $A_j^{(k)} = \pi_j^{(k)} - \pi_{j-1}^{(k)}$ et $B_j^{(k)} = (\nabla\pi_j)^{(k)}$. Puisque, par définition,

$$\pi_j^{(k+1)} = \pi_{j+k} + \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, \pi_j^{(\ell)}$$

et de même, pour $j > 1$,

$$\pi_{j-1}^{(k+1)} = \pi_{j-1+k} + \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, \pi_{j-1}^{(\ell)} .$$

on en déduit, en faisant la différence de ces deux égalités, que

$$A_j^{(k+1)} = \nabla\pi_{j+k} + \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, A_j^{(\ell)} , \text{ pour } j > 1 .$$

De même, pour la définition de $(\nabla\pi_j)^{(k)}$ qui est obtenue en appliquant l'opérateur $(\ )^{(k)}$ à la suite $\nabla\pi_j$, on voit que

$$B_j^{(k+1)} = \nabla\pi_{j+k} + \sum_{\ell=1}^{k} (\nabla\pi_{k-\ell+1})\, B_j^{(\ell)}$$

$$= \nabla\pi_{j+k} + \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, B_j^{(\ell)} - \sum_{\ell=1}^{k} \pi_{k-\ell}\, B_j^{(\ell)}$$

ou encore, puisque $\pi_0 = -I_n$,

$$B_j^{(k+1)} - \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, B_j^{(\ell)} = \nabla\pi_{j+k} + B_j^{(k)} - \sum_{\ell=1}^{k-1} \pi_{k-\ell}\, B_j^{(\ell)}$$

$$= \nabla\pi_{j+k} + B_j^{(k)} - \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, B_j^{(\ell-1)} ,$$

compte tenu de la convention $B_j^{(0)} = (\nabla\pi_j)^{(0)} = 0$. Ceci prouve qu'en notant $C_j^{(k+1)} = B_j^{(k+1)} - B_j^{(k)}$, on a

$$C_j^{(k+1)} = \nabla\pi_{j+k} + \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, C_j^{(\ell)} .$$

Ainsi, pour $j > 1$, les suites $A_j^{(k)}$ et $C_j^{(k)}$, $k \geq 1$, sont définies par la même formule de récurrence. Pour montrer qu'elles sont égales, il suffit donc de montrer que leurs valeurs initiales sont égales, ce qui est évident, car

$$A_j^{(1)} = \pi_j^{(1)} - \pi_{j-1}^{(1)} = \pi_j - \pi_{j-1} = \nabla\pi_j = (\nabla\pi_j)^{(1)} = (\nabla\pi_j)^{(1)} - (\nabla\pi_j)^{(0)} = C_j^{(1)} .$$

Cet argument peut être étendu au cas $j = 1$ grâce à la convention : $\pi_0^{(1)} = -I_n$ et $\pi_0^{(k)} = 0$ pour $k > 1$. En effet, cette convention rend valide l'égalité

$$\pi_{j-1}^{(k+1)} = \pi_{j-1+k} + \sum_{\ell=1}^{k} \pi_{k-\ell+1}\, \pi_{j-1}^{(\ell)}$$

même pour $j = 1$. On peut donc encore montrer de la même façon que les suites $A_1^{(k)}$ et $C_1^{(k)}$ sont définies par la même formule de récurrence. Leurs valeurs initiales coïncident puisque

$$A_1^{(1)} = \pi_1^{(1)} - \pi_0^{(1)} = \pi_1 + I_n = \nabla\pi_1 = C_1^{(1)} ,$$

ce qui achève la démonstration de (5.16) pour tout j. Q.E.D.

- Preuve de la proposition 24 : Supposons alors que $\pi_{XYj}^{(k)} = 0$ pour tout $j \in \mathbb{N}$ et pour tout $k = 1, 2, \ldots, h$. Il résulte de (5.16) que pour tout $k = 1, 2, \ldots, h$,

$$(\nabla\pi_j)_{XY}^{(k)} = (\nabla\pi_j)_{XY}^{(k-1)} + \pi_{XYj}^{(k)} - \pi_{XY,j-1}^{(k)} = (\nabla\pi_j)_{XY}^{(k-1)} = (\nabla\pi_j)_{XY}^{(0)} = 0 ,$$

puisque $\pi_{XYj}^{(k)} - \pi_{XY,j-1}^{(k)} = 0$ (y compris pour $j = 1$). Réciproquement, si $(\nabla\pi_j)_{XY}^{(k)} = 0$ pour tout j et pour tout $k = 1, 2, \ldots, h$, on a d'après (5.16),

$$\pi_{XYj}^{(k)} - \pi_{XY,j-1}^{(k)} = 0 , \text{ pour } k = 1, 2, \ldots, h .$$

Comme $\pi_{XY0}^{(k)} = 0$ en vertu de notre convention, on en déduit bien que $\pi_{XYj}^{(k)} = 0$ pour $k = 1, 2, \ldots, h$ et pour tout j. Q.E.D.

– **Preuve de la proposition 25 :** La démonstration procède en deux étapes.

Première étape : On vérifie que pour tout $t \in \mathbb{N}$ et pour tout $k = 1, 2, \ldots, h-1$:

$$[(\nabla^{\ell}\pi_t^*)^{(k)}]_{XZ}\,(\nabla^{\ell}\pi_t^*)^{(k)}{}_{XY} = 0\,,\ \text{pour } \ell = 0, 1, \ldots, d\,.$$

On sait, en effet, que $\pi^{*(k)}_{XYt} = 0$ pour $k = 1, 2, \ldots, h$, et donc aussi, en vertu de la proposition 24, que

$$[(\nabla^{\ell}\pi_t^*)^{(k)}]_{XY} = 0\,,\ \text{pour } k = 1, 2, \ldots, h \text{ et } \ell = 0, 1, \ldots, d.$$

Or, une réinterprétation algébrique immédiate de la proposition 8 prouve que

$$[(\nabla^{\ell}\pi_t^*)^{(k)}]_{XY} = 0 \text{ pour } k = 1, 2, \ldots, \bar{k} \Rightarrow [(\nabla^{\ell}\pi_t^*)^{(\bar{k}+1)}]_{XY} = [(\nabla^{\ell}\pi_t^*)^{(\bar{k})}]_{XZ}\,(\nabla^{\ell}\pi_t^*)_{ZY}\,.$$

On peut donc bien conclure que

$$[(\nabla^{\ell}\pi_t^*)^{(k)}]_{XZ}\,(\nabla^{\ell}\pi_t^*)_{ZY} = 0\,,\ \text{pour } k = 1, 2, \ldots, h-1\,.$$

Deuxième étape : Pour ℓ fixé ($0 \le \ell \le d$), on note Q_{ik} la propriété :

$$[(\nabla^{\ell+i}\pi_t^*)^{(k)}]_{XZ}\,(\nabla^{\ell}\pi_t^*)_{ZY} = 0\,,\ \forall t \in \mathbb{N}\,,$$

où $0 \le i \le d-\ell$. On note P_k la conjonction de propriétés Q_{ik} pour $i = 0, 1, \ldots, d-\ell$. On se propose de démontrer la propriété P_k pour $k = 1, 2, \ldots, h-1$ par récurrence sur k. Pour $k = 1$,

$$\nabla^{\ell+1}\pi_t^* = \nabla^{\ell}\pi_t^* - \nabla^{\ell}\pi_0^* = \nabla^{\ell}\pi_t^* + I_n$$

d'où

$$(\nabla^{\ell+1}\pi_t^*)_{XZ} = (\nabla^{\ell}\pi_t^*)_{XZ}\,.$$

Donc, pour $i = 0, 1, \ldots, d-\ell$,

$$(\nabla^{\ell+i}\pi_t^*)_{XZ}\,(\nabla^{\ell}\pi_t^*)_{ZY} = (\nabla^{\ell}\pi_t^*)_{XZ}\,(\nabla^{\ell}\pi_t^*)_{ZY} = 0$$

d'après la première étape. D'où la propriété P_1.

Pour $2 \le k \le h-1$, supposons maintenant que les propriétés $P_1, P_2, \ldots, P_{k-1}$ soient vérifiées. Il s'agit de démontrer la propriété P_k, c'est-à-dire la conjonction des propriétés Q_{ik} pour $i = 0, 1, \ldots, d-\ell$. D'après la première étape, la propriété Q_{0k} est vérifiée. On peut donc chercher à démontrer par récurrence sur i que toutes les propriétés Q_{ik}, $i = 0, 1, \ldots, d-\ell$ sont vérifiées. Supposons, pour cela, que les propriétés Q_{jk}, $j = 0, 1, \ldots, i$ ($i < d-\ell$) sont vérifiées (hypothèses de récurrence) et montrons que $Q_{i+1,k}$ l'est aussi. On sait par la proposition 23 que

$$(\nabla^{\ell+i+1}\pi_t^*)^{(k)} = (\nabla^{\ell+i+1}\pi_t^*)^{(k-1)} + (\nabla^{\ell+i}\pi_t^*)^{(k)} - (\nabla^{\ell+i}\pi_0^*)^{(k)}\,.$$

Par conséquent, comme $(\nabla^{\ell+i}\pi_0^*)^{(k)} = 0$, $(\nabla^{\ell+i+1}\pi_t^*)^{(k)}_{XZ}\,(\nabla^{\ell}\pi_t^*)_{ZY}$ est la somme de deux termes, $(\nabla^{\ell+i+1}\pi_t^*)^{(k-1)}_{XZ}\,(\nabla^{\ell}\pi_t^*)_{ZY}$ et $(\nabla^{\ell+i}\pi_t^*)^{(k)}_{XZ}(\nabla^{\ell}\pi_t^*)_{ZY}$, dont la nullité est assurée respectivement par les propriétés P_{k-1} et Q_{ik}. $Q_{i+1,k}$ est donc bien vérifiée, ce qui démontre par récurrence que les propriétés $P_1, P_2, \ldots, P_{h-1}$ sont vérifiées. On a donc, en particulier, les propriétés $Q_{d-\ell,k}$ pour $k = 1, 2, \ldots, h-1$:

$$[(\nabla^{d}\pi_t^*)^{(k)}]_{XZ}\,(\nabla^{\ell}\pi_t^*)_{ZY} = 0\,,$$

ce qui donne bien le résultat annoncé

$$\pi^{(k)}_{XZt}\,(\nabla^{\ell}\pi_t^*)_{ZY} = 0\,,\ \text{pour } k = 1, 2, \ldots, h-1\,. \qquad \text{Q.E.D.}$$

– **Preuve du théorème 2 :** D'après la proposition 24, la propriété (iii) est équivalente à

$$(\nabla\pi_j^*)^{(k)}_{XY} = 0\,,\ \text{pour } j \in \mathbb{N} \text{ et } k = 1, 2, \ldots, h\,.$$

En itérant d fois l'opérateur ∇, elle est aussi équivalente à

$$(\nabla^{d}\pi_j^*)^{(k)}_{XY} = 0\,,\ \text{pour } j \in \mathbb{N} \text{ et } k = 1, 2, \ldots, h\,.$$

En appliquant la proposition 21 à la représentation (5.15), on en déduit que la propriété (iii) est équivalente à la propriété (i). Il est par ailleurs évident que (iv) $\Rightarrow$ (ii) $\Rightarrow$ (i). En effet, (iv) signifie que $P[X(t+k) \mid I(t) + Y[0, t]]$, pour $k = 1, 2, \ldots, h$ et pour tout t est une transformation linéaire des composantes de $(X(\tau)', Z(\tau)')'$ pour $-d \le \tau \le t$, puisque $I(t) + Y[0, t] = X[0, t] + Z[0, t] = \bar{W}[-1] + Y[0, t]$ est l'espace vectoriel engendré par les composantes de $(X(\tau)', Y(\tau)', Z(\tau)')'$ pour $-d \leq \tau \leq t$. Si (iv) est vraie, $P[X(t+k) \mid I(t) + Y[0, t]]$ est donc aussi la projection de $X(t+k)$ (composante par composante) sur le sous-espace de $I(t) + Y[0, t]$ engendré par les composantes de $(X(\tau)', Z(\tau)')'$ pour $-d \leq \tau \leq t$, d'où la propriété (ii).

La propriété (ii), quant à elle, implique évidemment que $P[X(t+k) \mid I(t) + Y[0, t]]$ est aussi (pour tout $t \in \mathbb{N}_0$ et $k = 1, 2, \ldots, h$) la projection de $X(t+k)$, composante par composante, sur le sous-espace de $I(t) + Y[0, t]$ engendré par les composantes de $(X(\tau)', Z(\tau)')'$, $0 \leq \tau \leq t$, et de $\overline{W}_d(-1) = ((X(\tau)', Y(\tau)', Z(\tau)')' : \tau = -d, -d+1, \ldots, -1)'$. Donc, par définition de (i), (ii) implique (i).

L'équivalence des quatre propriétés (i), (ii), (iii) et (iv) sera donc démontrée si on prouve enfin que (iii) $\Rightarrow$ (iv). Supposons que

$$\pi^{*(k)}_{XYj} = 0\,, \text{ pour tout } j \in \mathbb{N} \text{ et } k = 1, 2, \ldots, h\,.$$

D'après la proposition 24, on en déduit que

$$(\nabla^{\ell} \pi^{*}_{j})^{(k)}_{XY} = 0\,, \text{ pour tout } j \in \mathbb{N},\ \ell = 0, 1, 2, \ldots, d, \text{ et } k = 1, 2, \ldots, h\,.$$

Or, d'après la proposition 20,

$$P[W(t+k) \mid I(t) + Y[0, t]] = \sum_{j=1}^{t+1} (\nabla^{d} \pi^{*}_{j})^{(k)}\, W(t+1-j) + C^{(k)}_{d}(t+1)\, \overline{W}_d(-1)\,.$$

Il est bien clair alors que pour $k = 1, 2, \ldots, h$, les valeurs de $Y(\tau)$, $-d \leq \tau \leq t$, n'interviennent pas dans le calcul des composantes numéros $1, 2, \ldots, m$ de $\sum_{j=1}^{t+1} (\nabla^{d} \pi^{*}_{j})^{(k)} X(t+1-j)$, où m est la dimension de $X(t)$. Il reste à montrer qu'elles n'interviennent pas non plus dans le calcul des composantes de mêmes numéros $1, 2, \ldots, m$ dans $C^{(k)}_{d}(t+1)\, \overline{W}_d(-1)$ pour $k = 1, 2, \ldots, h$.

Ceci est évident pour $k = 1$, puisque d'après la définition 6,

$$C^{(1)}_{d}(t+1)\, \overline{W}_d(-1) = C_d(t+1)\, \overline{W}_d(-1) = -\sum_{\ell=0}^{d-1} (\nabla^{\ell} \pi^{*}_{t})\, \nabla^{d-\ell-1}\, W(-1)$$

si bien que la nullité de $(\nabla^{\ell} \pi^{*}_{t})_{XY}$ pour tout $t \in \mathbb{N}$ et pour $\ell = 0, 1, 2, \ldots, d-1$ assure effectivement que les valeurs de Y n'interviennent pas dans le calcul des m premières composantes.

Le résultat va alors se déduire par récurrence sur k (où $1 \leq k \leq h-1$) en remarquant que, d'après la proposition 20 et la définition 6,

$$C^{(k+1)}_{d}(t+1)\, \overline{W}_d(-1) = C^{(k)}_{d}(t+2)\, \overline{W}_d(-1) + \pi^{(k)}_{1}\, C_d(t+1)\, \overline{W}_d(-1)$$

$$= C^{(k)}_{d}(t+2)\, \overline{W}_d(-1) - \sum_{\ell=0}^{d-1} \pi^{(k)}_{1} (\nabla^{\ell} \pi^{*}_{t})\, \nabla^{d-\ell-1} W(-1)\,.$$

Il s'agit donc, pour pouvoir appliquer l'argument de récurrence, de montrer que sous l'hypothèse (iii), le bloc $[\pi^{(k)}_{1}\, \nabla^{\ell} \pi^{*}_{t}]_{XY}$ de la matrice $\pi^{(k)}_{1}\, \nabla^{\ell} \pi^{*}_{t}$ est nul pour $\ell = 0, 1, \ldots, d-1$ et pour $k = 1, 2, \ldots, h-1$. On sera bien alors assuré que si, pour tout t, Y n'intervient pas dans le calcul des premières composantes de $C^{(k)}_{d}(t+2)\, \overline{W}_d(-1)$, il n'intervient pas non plus dans le calcul des m premières composantes de $C^{(k+1)}_{d}(t+1)\, \overline{W}_d(-1)$ et donc, par récurrence, le résultat cherché.

Sous l'hypothèse (iii), Y n'intervient pas dans le calcul des m premières composantes des vecteurs $C^{(i)}_{d}(t+1)\, \overline{W}_d(-1)$ pour $i = 1, 2, \ldots, h$. Or

$$[\pi^{(k)}_{1}\, \nabla^{\ell} \pi^{*}_{t}]_{XY} = \pi^{(k)}_{XX1} (\nabla^{\ell} \pi^{*}_{t})_{XY} + \pi^{(k)}_{XY1} (\nabla^{\ell} \pi^{*}_{t})_{YY} + \pi^{(k)}_{XZ1} (\nabla^{\ell} \pi^{*}_{t})_{ZY}\,,$$

et, d'après l'hypothèse (iii) ainsi que les propositions 24 et 25, nous avons

$$(\nabla^{\ell} \pi^{*}_{t})_{XY} = \pi^{(k)}_{XY1} = \pi^{(k)}_{XZ1} (\nabla^{\ell} \pi^{*}_{t})_{ZY} = 0$$

pour $k = 1, 2, \ldots, h$ et $\ell = 0, 1, \ldots, d$. Par conséquent, $[\pi^{(k)}_{1}\, \nabla^{\ell} \pi^{*}_{t}]_{XY} = 0$, pour $k = 1, \ldots, h$ et $\ell = 0, 1, \ldots, d$, ce qui complète la preuve du théorème. Q.E.D.

BIBLIOGRAPHIE

Ahn, S.K. et G.C. Reinsel (1988), « Nested Reduced-Rank Autoregressive Models for Multiple Time Series », *Journal of the American Statistical Association* 83, 849-856.

Boudjellaba, H., J.-M. Dufour et R. Roy (1992a), « Testing Causality Between Two Vectors in Multivariate ARMA Models », *Journal of the American Statistical Association* 87, 1082-1090.

Boudjellaba, H., J.-M. Dufour et R. Roy (1992b), « Simplified Conditions for Non-Causality Between Vectors in Multivariate ARMA Models », document de travail, C.R.D.E., Université de Montréal.

Dufour, J.-M. et D. Tessier (1991), « On the Relationship Between Impulse Response Analysis, Innovation Accounting and Granger Causality », document de travail, C.R.D.E., Université de Montréal.

Dufour, J.-M. et D. Tessier (1992), « Parsimonious Autoregressive Conditions for Noncausality in Multivariate ARMA Models », document de travail, C.R.D.E., Université de Montréal.

Florens, J.-P. et M. Mouchart (1985), « A Linear Theory for Non-Causality », *Econometrica* 53, 157-175.

Geweke, J. (1984), « Inference and Causality in Economic Time Series Models », Z. Griliches et M. Intrilligator (éd.), ***Handbook of Econometrics*** II, Amsterdam, North-Holland, 1102-1144.

Gouriéroux, C. et A. Monfort (1990), « Séries temporelles et modèles dynamiques », *Economica*, Paris.

Granger, C.W.J. (1969), « Investigating Causal Relations by Econometric Models and Cross-Spectral Methods », *Econometrica* 37, 424-438.

Hsiao, P. (1982), « Autoregressive Modeling and Causal Ordering of Economic Variables », *Journal of Economic Dynamics and Control* 4, 243-259.

Hosoya, Y. (1977), « On the Granger Condition for Non-Causality », *Econometrica* 45, 1735-1736.

Lütkepohl, H. (1990), « Testing for Causation Between Two Variables in Higher Dimensional VAR Models », texte présenté au « 6th World Congress of the Econometric Society », Barcelone, août 1990.

Lütkepohl, H. (1991), « Introduction to Multiple Time Series Analysis », Springer-Verlag, Berlin.

Newbold, P. (1982), « Causality Testing in Economics », in : O.D. Anderson (éd.), *Time Series Analysis : Theory and Practice* 1, Amsterdam, North-Holland, 701-716.

Pierce, D.A. et L.D. Haugh (1977), « Causality in Temporal Systems : Characterizations and Survey », *Journal of Econometrics* 5, 265-293.

Renault, E. et A. Szafarz (1991), « True Versus Spurious Instantaneous Causality », cahier de recherche no 9103, Centre d'économie mathématique et d'économétrie, Université Libre de Bruxelles.

Sims, C. (1980), « Macroeconomics and Reality », *Econometrica* 48, 1-48.

Université de Montréal
Département de sciences économiques

Cahiers de recherche (Discussion Papers)
1992 à aujourd'hui (1992 to date)

9201 : Dionne, Georges and Robert Gagné, "Measuring Technical Change and Productivity Growth with Varying Output Qualities and Incomplete Panel Data", 32 pages.
9202 : Beaudry, Paul and Michel Poitevin, "The Commitment Value of Contracts Under Dynamic Renegotiation", 30 pages.
9203 : Dionne, Georges et Christian Gollier, "Simple Increases in Risk and their Comparative Statics for Portfolio Management", 22 pages.
9204 : Fortin, Nicole M., "Allocation Inflexibilities, Female Labor Supply and Housing Assets Accumulation : Are Women Working to Pay the Mortgage?", 42 pages.
9205 : Beaudry, Paul, Marcel Boyer et Michel Poitevin, "Le rôle du collatéral dans le report des investissements en présence d'asymétries d'information", 20 pages.
9206 : Brenner, Reuven, Marcel G. Dagenais et Claude Montmarquette, "The Declining Saving Rate : An Overlooked Explanation", 34 pages.
9207 : Tremblay, Rodrigue, "L'impact fiscal statique et dynamique de l'accession du Québec au statut de pays souverain", 33 pages.
9208 : Mercenier, Jean, "Completing the European Internal Market : A General Equilibrium Evaluation Under Alternative Market Structure Assumptions", 38 pages.
9209 : Martin, Fernand and Sonia Granzer, "Notes of Methods of Assessing the Value of the Damages to Forests Caused by Air Pollution", 21 pages.
9210 : Dionne, Georges et Robert Gagné, "Rendements d'échelle, progrès technique et croissance de la productivité dans les industries québécoise et ontarienne de transport par camion, 1981-1988", 28 pages.
9211 : Sprumont, Yves, "Continuous Strategyproof Mechanisms for Sharing Private Goods", 30 pages.
9212 : Tremblay, Rodrigue, "L'émergence d'un bloc économique et commercial nord-américain : la compétitivité de l'économie canadienne et la politique du taux de change", 37 pages.
9213 : Hollander, Abraham, "Restricting Intra-industry Quota Transfers in Agriculture: Who Gains, Who Loses?", 10 pages.
9214 : Mandel, Benedikt, Marc Gaudry and Werner Rothengatter, "Linear or Nonlinear Utility Functions in Logit Models? The Impact of German High Speed Rail Demand Forecasts", 17 pages.
9215 : Ghysels, Éric, "Christmas, Spring and the Dawning of Economic Recovery", 26 pages.
9216 : Canova, Fabio et Éric Ghysels, "Changes in Seasonal Patterns : Are They Cyclical", 38 pages.

9217 : Campbell, Bryan et Éric Ghysels, "Is the Outcome of the Federal Budget Process Unbiaised and Efficient? A Nonparametric Assessment", 34 pages.

9218 : Boismenu, Gérard, Nicolas Gravel et Jean-Guy Loranger, "Régime d'accumulation et approche de la régulation : un modèle à équations simultanées", 27 pages.

9219 : Dionne, Georges and Pascale Viala, "Optimal Design of Financial Contracts and Moral Hazard", 60 pages.

9220 : Desruelle, Dominique, Gérard Gaudet et Yves Richelle, "Complementarity, Coordination and Compatibility : An Analysis of the Economics of Systems", 52 pages.

9221 : Desruelle, Dominique et Yves Richelle, "The Investment Dynamics of a Duopoly: The Relative Importance of a Head Start", 34 pages.

9222 : Mercenier, Jean, "Can '1992' Reduce Unemployment in Europe? On Welfare and Employment Effects of Europe's Move to a Single Market", 36 pages.

9223 : Dufour, Jean-Marie, Éric Ghysels et Alastair Hall, "Generalized Predictive Tests and Structural Change Analysis in Econometrics", 46 pages.

9224 : Dufour, Jean-Marie et Marc Hallin, "Improved Eaton Bounds for Linear Combinations of Bounded Random Variables, with Statistical Applications", 25 pages.

9225 : Crampes, Claude et Abraham Hollander, "How Many Karats is Gold : Welfare Effects of Easing a Denomination Standard, 20 pages.

9226 : Bonomo, Marco et René Garcia, "Indexation, Staggering and Disinflation", 32 pages.

9227 : Dudley, Leonard et Jacques Robert, "A Non-Cooperative Model of Alliances and Warfare", 26 pages.

9228 : Arcand, Jean-Louis, "Structural Adjustment and the Organization of Agricultural Credit in Egypt", 29 pages.

9229 : Arcand, Jean-Louis, "Supply Response and Marketing Board Policy : The Case of Egyptian Cotton", 38 pages.

9230 : Cadot, Olivier et Dominique Desruelle, "R & D : Who Does the R, Who Does the D?", 18 pages.

9231 : Kollmann, Robert, "Incomplete Asset Markets and International Business Cycles", 36 pages.

9232 : Kollmann, Robert, "Consumptions, Real Exchange Rates and the Structure of International Asset Markets", 33 pages.

9233 : Arcand, Jean-Louis, "Growth and Social Custom", 37 pages.

9234 : Desruelle, Dominique, "Infant Industries and Imperfect Capital Markets : The Case Against Tariff Protection", 30 pages.

9235 : Mercenier, Jean et Nicolas Schmitt, "Sunk Costs, Free-Entry Equilibrium and Trade Liberalization in Applied General Equilibrium : Implications for 'Europe 1992'"42 pages.

9236 : Boudjellaba, Hafida, Jean-Marie Dufour et Roch Roy, "Simplified Conditions for Non-Causality between Vectors in Multivariate Arma Models", 24 pages.

9237 : Ghysels, Eric, Hahn S. Lee et Pierre L. Siklos, "On the (Mis)specification of Seasonality and its Consequences : An Empirical Investigation with U.S. Data", 45 pages.

9301 : Mercenier, Jean, "Nonuniqueness of Solutions in Applied General-Equilibrium Models with Scale Economies and Imperfect Competition : A Theoretical *Curiosum*?", 26 pages.

9302 : Lemieux, Thomas, "Unions and Wage Inequality in Canada and in the United States", 66 pages.

9303 : Lemieux, Thomas, "Estimating the Effects of Unions on Wage Inequality in a Two-Sector Model with Comparative Advantage and Non-Random Selection", 54 pages.

9304 : Harchaoui, Tarek H., "Time-Varying Risks and Returns : Evidence from Mining Industries Data", 20 pages.

9305 : Lévy-Garboua, Louis et Claude Montmarquette, "Une étude économétrique de la demande de théâtre sur données individuelles", 40 pages.

9306 : Montmarquette, Claude, Rachel Houle et Sophie Mahseredjian, "The Determinants of University Dropouts : A Longitudinal Analysis", 17 pages.

9307 : Gaudry, Marc, Benedikt Mandel et Werner Rothengatter, "A Disaggregate Box-Cox Logit Mode Choice Model of Intercity Passenger Travel in Germany", 17 pages.

9308 : Fortin, Nicole M., "Borrowing Constraints and Female Labor Supply : Nonparametric and Parametric Evidence of the Impact of Mortgage Lending Rules", 38 pages.

9309 : Dionne, Georges, Robert Gagné, François Gagnon et Charles Vanasse, "Debt, Moral Hazard and Airline Safety : an Empirical Evidence", 34 pages.

9310 : Dionne, Georges, Anne Gibbens et Pierre St-Michel, "An Economic Analysis of Insurance Fraud", 40 pages.

9311 : Gaudry, Marc, "Asymmetric Shape and Variable Tail Thickness in Multinomial Probabilistic Responses to Significant Transport Service Level Changes", 26 pages.

9312 : Laferrière, Richard et Marc Gaudry, "Testing the Linear Inverse Power Transformation Logit Mode Choice Model", 29 pages.

9313 : Kollmann, Robert, "Fiscal Policy, Technology Shocks and the US Trade Balance Deficit", 38 pages.

9314 : Ghysels, Eric, "A Time Series Model With Periodic Stochastic Regime Switching", 54 pages.

9315 : Allard, Marie, Camille Bronsard et Lise Salvas-Bronsard, "C^*-Conjugate Expectations and Duality", 22 pages.

9316 : Dudley, Leonard et Claude Montmarquette, "Government Size and Economic Convergence", 28 pages.

9317 : Bronsard, Camille, "L'histoire de l'économie mathématique racontée à Juliette", 17 pages.

9318 : Tremblay, Rodrigue, "The Quest for Competitiveness and Export-Led Growth", 16 pages.

9319 : Proulx, Pierre-Paul, "L'ALÉNA", 12 pages.

9320 : Proulx, Pierre-Paul, "Le Québec dans l'ALÉNA", 28 pages.
9321 : Dionne, Georges, Denise Desjardins, Claire Laberge-Nadeau et Urs Magg, "Medical Conditions, Risk Exposure and Truck Drivers' Accidents : an Analysis with Count Data Regression Models, 20 pages.
9322 : À paraître.
9323 : Dufour, Jean-Marie et David Tessier, "On the Relationship Between Impulse Response Analysis, Innovation Accounting and Granger Causality", 12 pages.
9324 : Dufour, Jean-Marie et Eric Renault, "Causalités à court et à long terme dans les modèles VAR et ARIMA multivariés", 68 pages.