

# Estimation spectrale Paramétrique ou pas

Introduction

[jean-christophe.cexus@ensta-bretagne.fr](mailto:jean-christophe.cexus@ensta-bretagne.fr)

**ENSTA-Bretagne**

**Laboratoire STIC/REMS // LabSTICC UMR 6286**

“ Spectral Estimation is ... an Art ”

# A propos du cours

Petre Stoica

- Le cours d'*Estimation Spectrale* décrit des méthodes de base d'estimation et d'analyse du spectre.
- Le cours peut être un bon point de départ pour de nombreuses autres techniques d'estimation et de problèmes avancés (Filtrage de Wiener, Kalman ...).
- Le livre de base :
  - **Petre Stoica and Randolph Moses, « Spectral analysis of signals », Prentice Hall, 2005**
  - Petre Stoica and Randolph Moses, « Introduction to spectral analysis », Prentice Hall, 1997
- Les autres bonnes sources sont:
  - **S. M. Kay, « Modern spectral Estimation », Prentice Hall, 1988.**
  - S. L. Marple Jr., « Digital spectral analysis: with applications », Prentice Hall, 1987.
  - F. Castanié, « Analyse spectrale / sous la direction de Francis Castanié », Hermès science, Lavoisier, IC2, 2003.
  - J. G. Proakis et D. G. Manolakis, « Digital signal processing: Principles, algorithms, and applications », Prentice Hall, 1996.

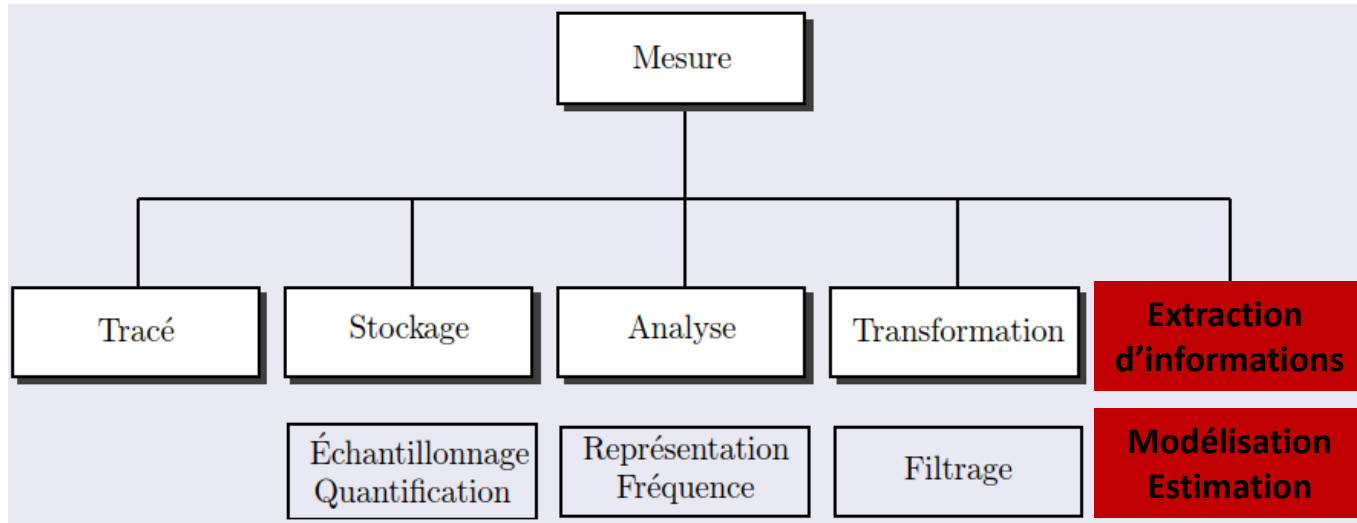
# Plan du cours

- **Introductions**
- **Analyse spectrale classique**
  - Focus sur quelques notions : stationnaire, aléatoire, ergodique
  - Cadre déterministe et aléatoire
  - Périodogramme & Corrélogramme & Extensions
  - **Mise en œuvre implémentation sous Matlab (BE)**
- **Vers les Modèles rationnels linéaires**
  - Retour (ou pas) sur les SLI
  - Décomposition spectrale : Modèles ARMA, AR, MA ...
  - Relation avec l'analyse spectrale classique : le modèle de Yule-Walker
  - **Estimation des paramètres des modèles** : inversion directe, Moindres carrés *Levinson*, ...

## Conclusions

# Introduction

- Intérêt
  - L'analyse spectrale nous permet d'avoir une « estimation » du spectre de puissance d'un signal à partir d'un nombre fini d'observations



- Applications : sonar, radar, médecine, économie, optique, ...

# Introduction

- **Objectif**

- Représentation d'un signal dans le domaine spectral
- **Déterminer la distribution de la puissance du signal en fréquences :**  
densité spectrale

- **Contexte**

- En analyse spectrale, on s'intéresse essentiellement à l'estimation du contenu spectral de signaux aléatoires :
  - Une seule réalisation de disponible (avec souvent peu d'échantillons)
  - Problème d'estimation
- Présentation des outils d'estimation du spectre d'un signal aléatoire.

# Introduction

## Un peu d'histoire

- Newton (1671) décompose la lumière et utilise le terme « *spectrum* » (observée à travers un prisme).
  - Etymologie de *specio* (« regarder ») avec le suffixe –trum
  - Signifie **spectre**, simulacre émis par les objets ou apparition fantomatique.
  - Son adjectif est spectral.
  - Le terme souvent utilisé et préféré étant *Spectral Analysis*.
- Bernoulli (1738) résolution de l'équation d'onde

$$u(x, t) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{a}\right) \left[ a_n \cos\left(\frac{n\pi ct}{a}\right) + b_n \sin\left(\frac{n\pi ct}{a}\right) \right]$$



- Euler (1755) démontre que les coefficients  $a_n$  et  $b_n$  sont données par les séries de Fourier

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} f(x) \cos n\omega x \, dx \quad b_n = \frac{2}{T} \int_{-T/2}^{T/2} f(x) \sin n\omega x \, dx \quad \text{si } n \geq 1$$

# Introduction

- Fourier (1822) affirme que l'on peut décomposer un signal périodique (l'analyse harmonique consistant alors à déterminer  $a_n$  et  $b_n$ )

$$f(x) = \sum_{n=0}^{+\infty} a_n \cos(n\omega x) + b_n \sin(n\omega x), \quad a_n, b_n \in \mathbb{R}$$



- La démonstration sera faite par Dirichlet en 1929.
- Au milieu du XIXe siècle, l'analyse harmonique est utilisée pour étudier le son (la musique), la météo, les taches solaires, le débit des cours d'eau ...
- Des analyseurs harmoniques mécaniques vont voir le jour dans les années 1900 permettant d'obtenir jusqu'à 150 coefficients du développement de Fourier ! (Lord Kelvin 1878).
- Permet d'analyser essentiellement des périodicités évidentes peu ou pas bruitées.
- Les périodicités cachées (signal périodique non évident) ne sont pas résolues.
- Schuster développe un estimateur spectral appelé *periodogram* permettant de trouver des périodicités cachées.

" It is convenient to have a word for some representation of a variable quantity which shall correspond to the 'spectrum' of a luminous radiation. I propose the word **periodogram**, and define it more particularly in the following way. "

# Introduction

- **Yule** (1927) a recours à l'(auto)régression linéaire pour trouver des périodicités dans des séries chronologiques. Walker (1931) aussi. C'est la raison pour laquelle **les équations normales** sont parfois appelées les **équations de Yule-Walker**.
- **Wiener** (1930) décrit la densité spectrale de puissance (DSP) des signaux stationnaires (Théorème de Wiener-Khintchine).
- **Wold** (1938) introduit les termes **Moving Average (MA)** et **AutoRegression (AR)** pour la prediction linéaire.
- **Levinson** (1947) propose une solution rapide pour résoudre les équations normales (matrice de Toeplitz).
- **Cooley et Tukey** (1965) propose l'algorithme *Fast Fourier Transform*.



# Introduction

- **Avec Burg** (1967), l'estimation spectrale haute résolution débute :
  - Méthode des moindres carrées non-linéaires (Nonlinear least squares) pour les spectres continus (méthode développée initialement par Gauss).
  - Début des années 60, apparition des premiers algorithmes rapides pour les moindres carrés
  - Méthode de Capon / Minimum de variance (1969).
  - Méthode de Pisarenko (1973).
  - Méthode MUSIC par Schmidt et Bienvenu (1979).
  - Méthode par Banc de filtres (1982),
  - Méthode Min-Norm par Kumaresan et Tufts (1983).
  - Méthode ESPRIT par Paulraj, Roy et Kailath (1986).
  - ...

# Introduction

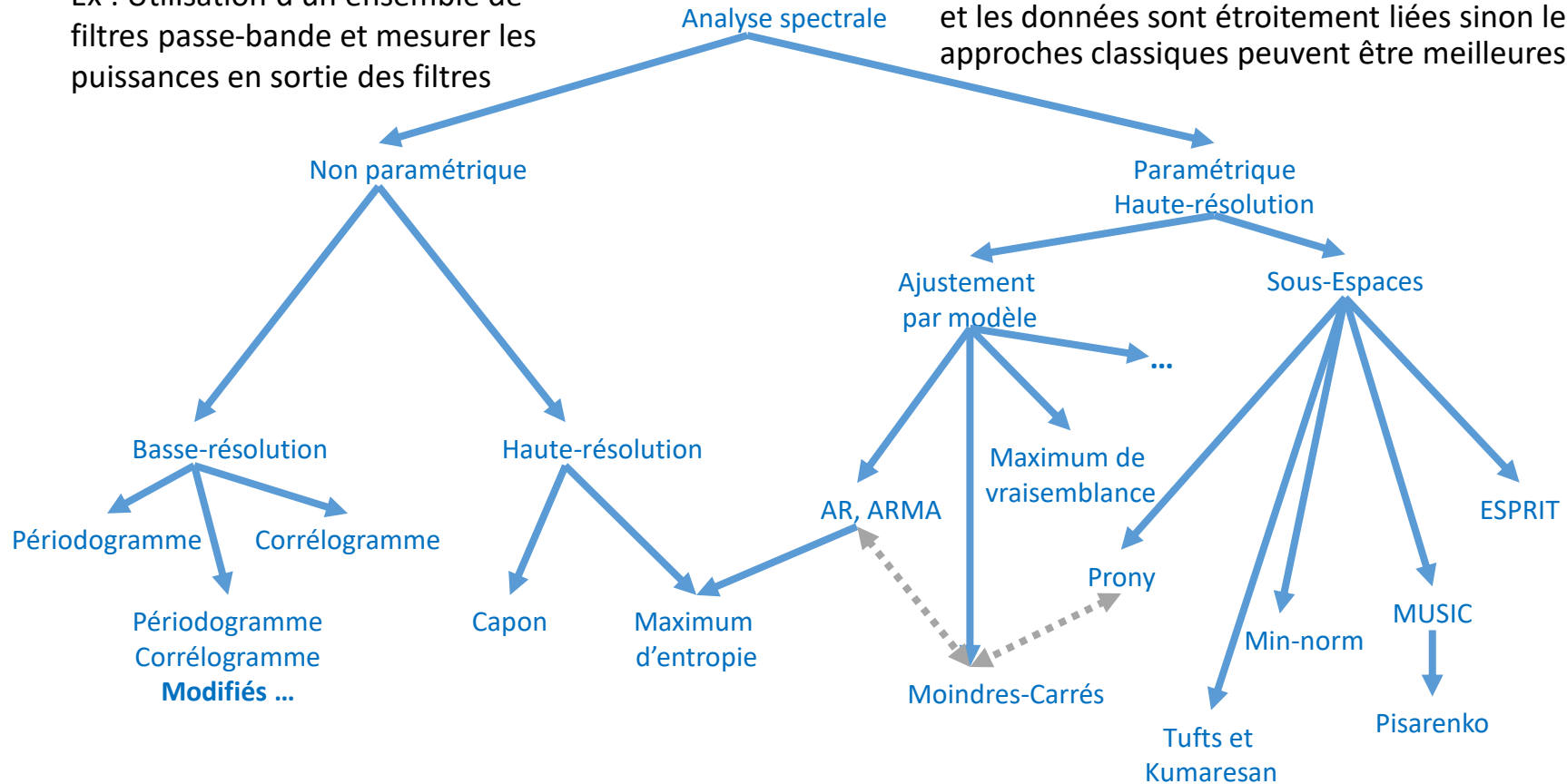
## Approche classique (non paramétrique)

Ex : Utilisation d'un ensemble de filtres passe-bande et mesurer les puissances en sortie des filtres

## Approche moderne (paramétrique)

Ex : Modéliser sous la forme de somme de sinusoides amorties et en estimer leurs paramètres

Peut permettre de meilleures estimations si le modèle et les données sont étroitement liées sinon les approches classiques peuvent être meilleures.



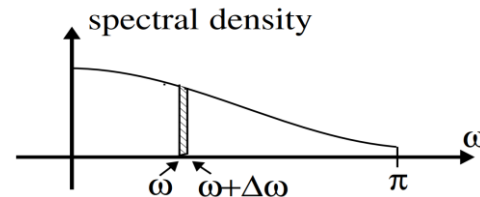
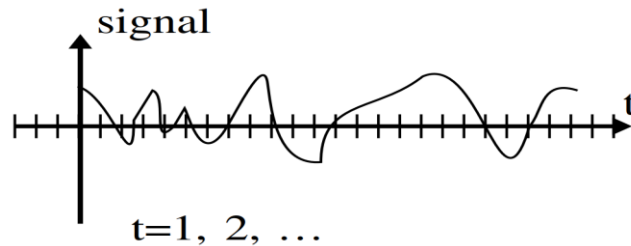
# Introduction

- **Quelques applications**

- Parole :
  - Modélisation du processus de la parole (extraction des formants pour la reconnaissance)
  - Compression et codage
- Sonar, Radar :
  - Localisation de source avec des antennes linéaires
  - Imagerie Radar (SAR ISAR) et extraction de paramètres d'une cible
- Electromagnétique:
  - Fréquence de résonance d'une cavité
- Communications:
  - Codage des transmissions par étalement du spectre (Code Division Multiple Access CDMA)
- Et encore : médecine, économie, optique, ...

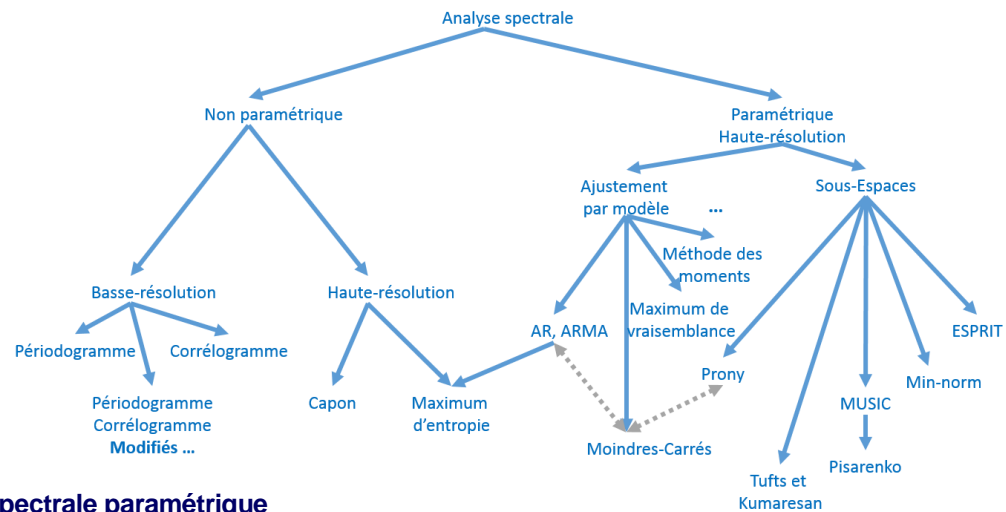
# Introduction

- Hypothèses générales
  - Signaux à valeurs réelles
  - **Signaux à temps discrets : N échantillons disponibles**
- L'estimation spectrale suppose **la stationnarité du processus** (même à court terme), sinon le spectre en tant que tel n'est pas définie
- On suppose aussi que **le processus est ergodique** ce qui nous permet d'estimer les moyennes d'ensemble par les moyennes temporelles



$$f = \omega / 2\pi$$

# Focus sur aléatoire, stationnaire, ergodique



# Processus aléatoires à temps discret

- **Processus et modèles aléatoires discrets.**

- **Processus stochastique ou processus aléatoire**

On parle de processus stochastique ou processus aléatoire pour décrire l'évolution au cours du temps d'un phénomène statistique selon des lois de probabilité.

Due à l'incertitude, il n'est pas possible de décrire exactement comment le processus va évoluer dans le temps avant de conduire l'expérience.

Une réalisation particulière d'un processus stochastique à temps discret est une suite temporelle. Par exemple, la séquence  $x(n)$ ,  $x(n-1)$ ,  $\dots$ ,  $x(n-L+1)$  représente une suite temporelle qui est constituée d'une observation au présent  $x(n)$  et de  $L-1$  observations du passé.

Dans un processus déterministe, il existe une formule pour le définir exactement.

- **Caractérisation partielle d'un processus stochastique à temps discret**

La densité de probabilité conjointe d'un ensemble arbitraire d'observations d'un processus aléatoire caractérise complètement ce processus. Cependant, cette densité de probabilité n'est pas facile à obtenir en pratique.

On se contentera donc d'une caractérisation partielle avec des moments jusqu'à l'ordre 2.

Un processus est dit du second ordre s'il possède un moment du premier ordre appelé la moyenne et s'il possède un moment du second ordre appelé l'autocorrélation.

# Processus aléatoires à temps discret

- Processus et modèles aléatoires discrets.

- Moyenne, autocorrélation et stationnarité

Soit un processus stochastique **discret** représenté par la série temporelle  $X(n)$ ,  $X(n-1)$ ,  $\dots$ ,  $X(n-M)$ , on définit la **moyenne** par :

$$\mu_X(n) = E\{X(n)\}$$

où  $E\{\cdot\}$  est l'opérateur espérance mathématique (moyenne d'ensemble). De même, la **fonction d'autocorrélation (ACF)** prend la forme :

$$r(n, n-k) = E\{X(n)X^*(n-k)\}, \quad k = 0, \pm 1, \pm 2, \dots$$

où l'astérisque représente la conjugaison complexe. La **fonction d'autocovariance** s'écrit :

$$c(n, n-k) = E\{[X(n) - \mu_X(n)][X(n-k) - \mu_X(n-k)]^*\}, \quad k = 0, \pm 1, \pm 2, \dots$$

$$c(n, n-k) = r(n, n-k) - \mu_X(n)\mu_X^*(n-k)$$

Un processus est dit **strictement stationnaire** si tous ses moments sont indépendants du temps. Un processus est dit faiblement stationnaire, ou **stationnaire au sens large (SSL)**, si

$$\mu_X(n) = \mu \quad \forall n$$

$$r(n, n-k) = r(k) \quad \forall n$$

On notera au passage que  **$r(0)$**  représente la *valeur quadratique moyenne* ou *puissance* de  $X(n)$ , tandis que  **$c(0) = \sigma^2$**  représente la **variance**.

$$r(0) \geq |r(k)| \quad r(k) = r^*(-k)$$

# Processus aléatoires à temps discret

- **Processus et modèles aléatoires discrets.**

- **La matrice d'autocorrélation.**

Définissons un *vecteur d'observation* (qui est un vecteur aléatoire !)  $\mathbf{X}(n)$  SSL à temps discret tel que :

$$\mathbf{X}^T(n) = [X(n), X(n-1), \dots, X(n-M+1)]$$

On définit alors la **matrice d'autocorrélation** par :

$$\mathbf{R} = E \{ \mathbf{X}(n) \mathbf{X}^H(n) \}$$

où  $\mathbf{X}^H(n)$  représente la transposée hermitienne de  $\mathbf{X}(n)$ , c'est-à-dire le vecteur transposé conjugué.

En détaillant la **matrice d'autocorrélation**, on obtient immédiatement :

$$\mathbf{R} = \begin{bmatrix} r(0) & r(-1) & \dots & r(-M+1) \\ r(1) & r(0) & \dots & r(-M+2) \\ \vdots & \vdots & \ddots & \vdots \\ r(M-1) & r(M-2) & \dots & r(0) \end{bmatrix} \quad r(k) = r^*(-k)$$

- *La matrice d'autocorrélation d'un processus stochastique discret stationnaire est **hermitienne**.*

Une matrice est dite hermitienne si elle est égale à sa transposée hermitienne, i.e.  $\mathbf{R} = \mathbf{R}^H$

Cette propriété découle directement de la définition de la matrice de corrélation. Dans le cas d'un processus à valeurs réelles, la matrice  $\mathbf{R}$  est simplement symétrique.

- *La matrice d'autocorrélation d'un processus stationnaire discret est une **matrice de Toeplitz**.*

Une matrice carrée est dite Toeplitz si tous les éléments d'une même diagonale ou sous-diagonale sont égaux. Cette propriété est directement liée à la propriété de stationnarité (au sens large) du processus. Cette propriété est importante, car elle permet dans bien des cas de simplifier les calculs algébriques.



# Processus aléatoires à temps discret

- Processus et modèles aléatoires discrets.

- Ergodicité

Dans le cas d'un signal réel dont on a procédé à une seule acquisition lors d'une expérience  $\xi_0$ , le problème est de déterminer ses différentes caractéristiques statistiques :  $\mu$ ,  $\sigma^2$ , ... Une idée est que si un signal aléatoire est stationnaire, il est possible d'avoir une estimation de la moyenne et de la variance en intégrant sur le temps  $t$  à partir des valeurs  $x(t, \xi_0)$  correspondant à cette seule acquisition.

$$m_x \approx \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t, \xi_0) dt \quad \sigma_x^2 \approx \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} (x(t, \xi_0) - m_x)^2 dt.$$

Un processus aléatoire est ergodique au sens strict, si tous les moments statistiques sont égaux aux moments temporels.

Un processus aléatoire est ergodique au sens large si il y a égalité des moyennes statistiques et temporelles ainsi que des fonctions d'autocorrélation.

$$m = E\{x(n)\} = \lim_{N \rightarrow \infty} \left[ \frac{1}{2N+1} \sum_{n=-N}^N x(n) \right]$$

$$\begin{aligned} r(k) &= E\{x(n)x^*(n-k)\} \\ &= \lim_{N \rightarrow \infty} \left[ \frac{1}{2N+1} \sum_{n=-N}^N x(n)x^*(n-k) \right] \end{aligned}$$

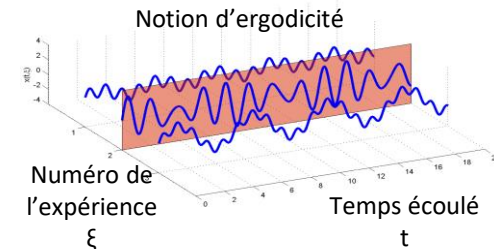
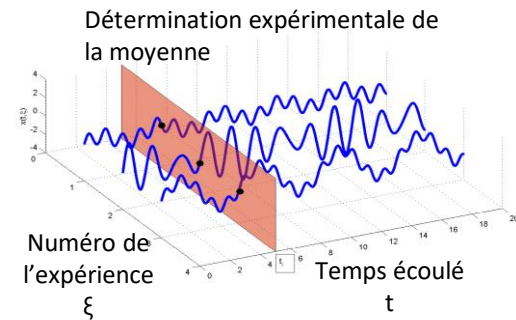
- Bruit blanc (un signal aléatoire stationnaire et ergodique)

On appelle bruit blanc à temps discret un processus aléatoire SSL, centré, dont la DSP est constante sur tout l'axe des fréquences

$$\phi_w(f) = \sigma_w^2$$

Si on note  $\sigma^2$  la DSP d'un bruit blanc, une conséquence directe de la définition est que la fonction d'autocorrélation d'un bruit blanc est :

$$E\{w(n)w(n-k)\} = \begin{cases} \sigma_w^2 & \text{si } k = 0 \\ 0 & \text{si } k \neq 0 \end{cases}$$

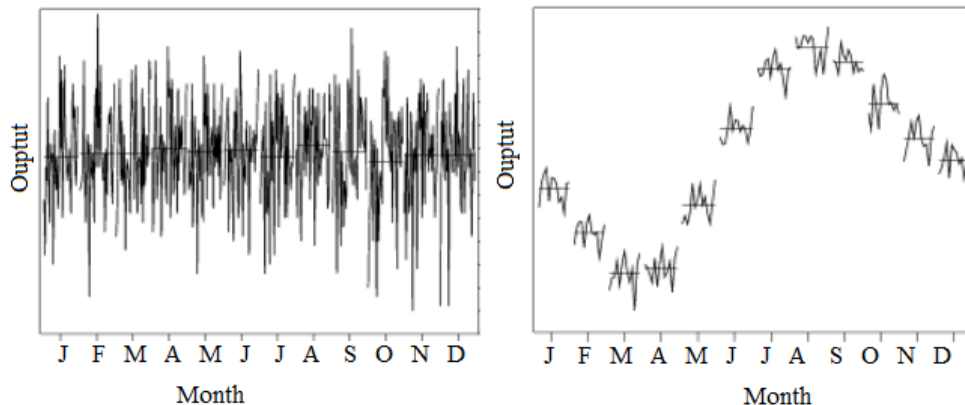


# Processus Stationnaire

“Experience with real-world data, however, soon convinces one that both stationarity and Gaussianity are fairy tales invented for the amusement of undergraduates.” Thomson (1994)

- **Retour sur la stationnarité**

Un signal est stationnaire si un décalage dans le temps ne modifie pas la forme de sa distribution. Les propriétés de base de la distribution telles que la moyenne, la variance et sa covariance (autocorrélation) sont constantes dans le temps.



Le signal de gauche est stationnaire, sans tendance évidente, tandis que le signal de droite montre une *'saisonnalité'* et n'est pas stationnaire.

- **La stationnarité est-elle importante ?**

La plupart des méthodes de prévision supposent qu'une distribution est stationnaire. Par exemple, l'autocovariance et l'autocorrélation reposent sur l'hypothèse de la stationnarité. Une absence de stationnarité peut entraîner des comportements inattendus ou bizarres en terme statistiques (distribution, variables corrélées, peu d'échantillons)

# Processus Stationnaire

“Experience with real-world data, however, soon convinces one that both stationarity and Gaussianity are fairy tales invented for the amusement of undergraduates.” Thomson (1994)

- **Retour sur la stationnarité**

Les modèles peuvent avoir différents types de stationnarité:

- **La stationnarité stricte** signifie que la distribution conjointe de tous les moments de tout degré (par exemple, les valeurs attendues, les variances, les moments du troisième ordre et les moments les plus élevés) dans le processus ne dépend jamais du temps. Cette définition est en pratique trop stricte pour être utilisée dans un modèle réel.
- **La stationnarité au premier ordre** ont des moyennes qui ne changent jamais avec le temps. Toute autre statistique (comme la variance) peut changer.
- **La stationnarité au second ordre (ou stationnaire au sens large (SSL))** ont une moyenne, une variance et une autocorrélation (autocovariance) constantes qui ne changent pas avec le temps. Les autres statistiques du système sont libres de changer au fil du temps. Cette version contrainte de la stricte stationnarité est très courante.
- **Les modèles à tendance stationnaire** fluctuant autour d'une tendance déterministe (la moyenne de la série). Ces tendances déterministes peuvent être linéaires ou quadratiques, mais l'amplitude (la hauteur d'une oscillation) des fluctuations n'augmente ni ne diminue à travers la série.
- **Les modèles stationnaires par différence** sont des modèles qui nécessitent une ou plusieurs différences pour devenir stationnaires (voir Transformation de modèles – logarithme, différences ... ).

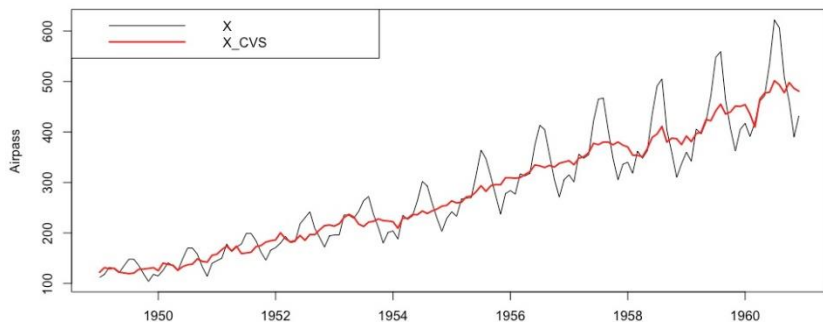
# Processus Stationnaire

“Experience with real-world data, however, soon convinces one that both stationarity and Gaussianity are fairy tales invented for the amusement of undergraduates.” Thomson (1994)

- **Retour sur la stationnarité**

Il peut être difficile de savoir si un modèle est stationnaire ou non. Généralement, Il est quasi impossible de savoir en regardant un graphique. Si la stationnarité d'un modèle n'est pas évidente, il existe des **test d'hypothèse** comme par exemple :

- Tests de racine unitaire (par exemple le Test de Dickey-Fuller augmenté ou le test de Zivot-Andrews),
- Test de Kwiatkowski–Phillips–Schmidt–Shin (KPSS) (en complément des tests de racine unitaire).
- Le test de Priestley-Subba Rao (PSR) ou test basé sur les ondelettes, tests moins fréquents basés sur l'analyse spectrale



Correction en variations saisonnières par régression linéaire

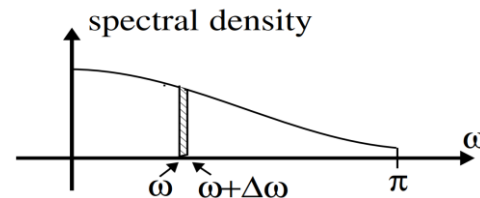
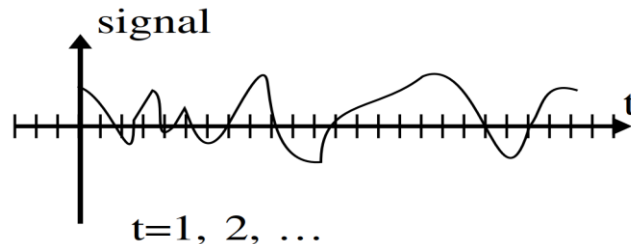
Nombre mensuel de passagers aériens, en milliers, de janvier 1949 à décembre 1960

Priestley, M. & Subba Rao, T. (1969) A Test for Non-Stationarity of Time-Series. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 31, No. 1, pp. 140-149.

Von Sachs, R. & Neumann, M. (2000). A Wavelet-Based Test for Stationarity. Journal of Time Series Analysis. Vol. 21, No. 5. September, pp. 597-613.

# Introduction

- Hypothèses générales
  - Signaux à valeurs réelles
  - **Signaux à temps discrets : N échantillons disponibles**
- L'estimation spectrale suppose **la stationnarité du processus** (même à court terme), sinon le spectre en tant que tel n'est pas définie
- On suppose aussi que **le processus est ergodique** ce qui nous permet d'estimer les moyennes d'ensemble par les moyennes temporelles



$$f = \omega / 2\pi$$

# Cadre déterministe

Soit  $\{y(t)\}_{t=-\infty}^{\infty}$  un signal déterministe à temps discret.

$$y(t) = y_c(t \cdot T_s)$$

Si  $\sum_{t=-\infty}^{\infty} |y(t)|^2 < \infty$  (signaux à énergie finie)

Alors la Transformée de Fourier à Temps Discret (TFTD) existe :

$$Y(\omega) = \sum_{t=-\infty}^{\infty} y(t) e^{-i\omega t} \quad \omega \in [-\pi, \pi]$$

Et sa Transformée Inverse :

$$y(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} Y(\omega) e^{i\omega t} d\omega$$

# Cadre déterministe

$$Y(\omega) = \sum_{t=-\infty}^{\infty} y(t)e^{-i\omega t}$$

La Densité Spectrale d'Energie (DSE) :

$$S(\omega) = |Y(\omega)|^2$$

Théorème de Parseval :

$$\sum_{t=-\infty}^{\infty} |y(t)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) d\omega$$

En définissant la '*fonction d'autocorrélation*' de  $y(t)$  :  $\rho(k) = \sum_{t=-\infty}^{\infty} y(t)y^*(t-k)$

Alors on a aussi :

$$S(\omega) = \sum_{k=-\infty}^{\infty} \rho(k)e^{-i\omega k}$$

# Cadre aléatoire

$$y(t) = y_c(t \cdot T_s)$$

Soit  $\{y(t)\}_{t=-\infty}^{\infty}$  une réalisation (trajectoire) issu d'un processus aléatoire **SSL** à **temps discret**

Dans ce cas :

$$\sum_{t=-\infty}^{\infty} |y(t)|^2 = \infty$$

Mais :

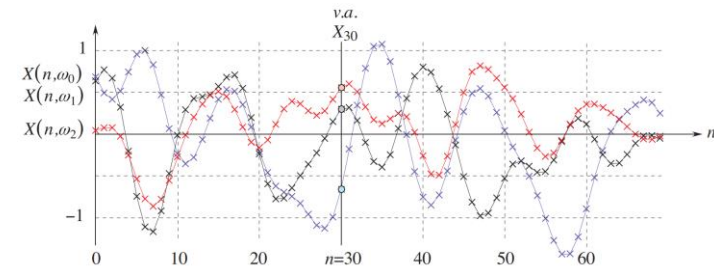
$$E\{|y(t)|^2\} < \infty \quad (\text{puissance moyenne finie})$$

Avec  $E\{. \}$  est l'Espérance sur l'ensemble des réalisations

Par la suite, pour simplifier les calculs, et s'en perdre en généralité, on supposera que les processus seront SSL de moyenne nulle

$$E\{y(t)\} = 0.$$

La contrainte  $\mu = 0$  n'est pas très forte puisqu'on passe d'un processus d'espérance  $\mu$  à un processus d'espérance nulle par simple translation.





# Cadre aléatoire

## 1ere définition

Densité spectrale de puissance (moyenne) est :

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}$$

avec la fonction d'autocorrelation :

$$r(k) = E \{y(t)y^*(t - k)\}$$

([ACF – autocovariance function](#))

$$r(k) = r^*(-k)$$

$$r(0) \geq |r(k)|$$

Et son inverse

$$r(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\omega)e^{i\omega k} d\omega$$

$$\begin{aligned} r(0) &= E \{|y(t)|^2\} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\omega) d\omega \end{aligned}$$

# Cadre aléatoire

## 2eme définition

Densité spectrale de puissance (moyenne) est :

$$\phi(\omega) = \lim_{N \rightarrow \infty} E \left\{ \frac{1}{N} \left| \sum_{t=1}^N y(t) e^{-i\omega t} \right|^2 \right\}$$

A noter que :

$$\phi(\omega) = \lim_{N \rightarrow \infty} E \left\{ \frac{1}{N} |Y_N(\omega)|^2 \right\}$$

Avec

$$Y_N(\omega) = \sum_{t=1}^N y(t) e^{-i\omega t}$$

Cadre déterministe

$$Y(\omega) = \sum_{t=-\infty}^{\infty} y(t) e^{-i\omega t}$$

# Quelques propriétés

- Périodicité de  $2\pi$  :  $\phi(\omega) = \phi(\omega + 2\pi)$

ainsi, il est possible de restreindre notre étude sur :

$$\omega \in [-\pi, \pi] \iff f \in [-1/2, 1/2] \qquad f = \omega/2\pi$$

- C'est une densité de puissance :

$$\phi(\omega) \geq 0$$

- Si  $y(t)$  est réel :

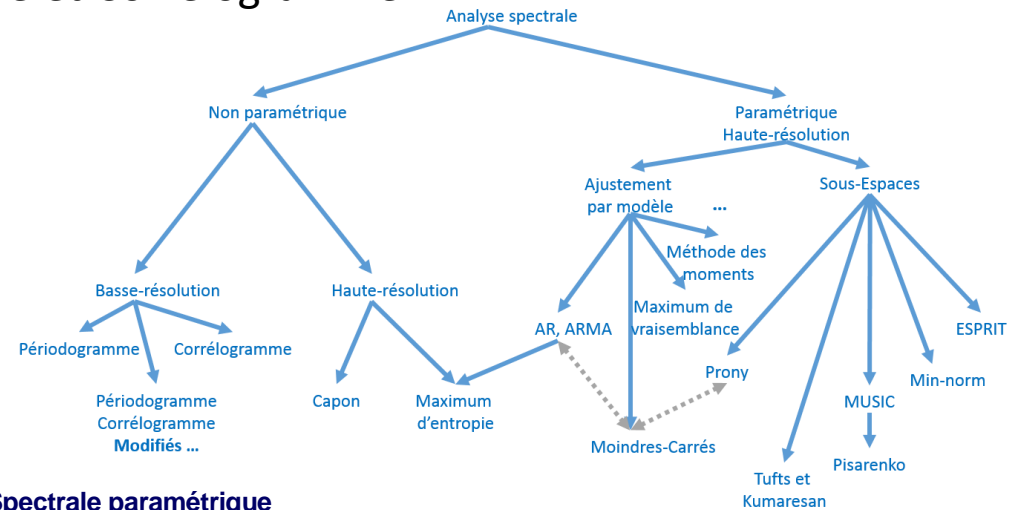
$$\phi(\omega) = \phi(-\omega), \quad \omega \in [-\pi, \pi] \qquad \text{spectre symétrique}$$

sinon

$$\phi(\omega) \neq \phi(-\omega), \quad \omega \in [-\pi, \pi]$$

# Estimation spectrale

## Périodogramme et Corrélogramme



# Périodogramme

- En se basant sur la 2ème définition de la DSP

$$\phi(\omega) = \lim_{N \rightarrow \infty} E \left\{ \frac{1}{N} \left| \sum_{t=1}^N y(t) e^{-i\omega t} \right|^2 \right\}$$

- Utilisation directe de  $\{y(t)\}_{t=1}^N$

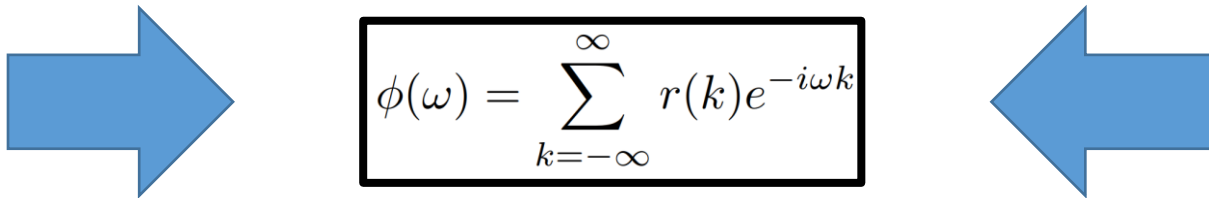
En faisant abstraction de la limite et de l'espérance :

$$\hat{\phi}_p(\omega) = \frac{1}{N} \left| \sum_{t=1}^N y(t) e^{-i\omega t} \right|^2$$

Estimateur naturel  
Proposé par A. Schuster 1898

# Corrélogramme

- En se basant sur la 1ère définition de la DSP


$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}$$

En tronquant la somme et en estimant la grandeur  $r(k) \sim \hat{r}(k)$  :

$$\hat{\phi}_c(\omega) = \sum_{k=-(N-1)}^{N-1} \hat{r}(k)e^{-i\omega k}$$

# Estimation de la covariance (ACF)

$$r(k) = E \{y(t)y^*(t-k)\}$$

- **Estimateur non-biaisé**

$$\hat{r}(k) = \frac{1}{N-k} \sum_{t=k+1}^N y(t)y^*(t-k), \quad 0 \leq k \leq N-1$$

- **Estimateur biaisé**

$$\hat{r}(k) = \frac{1}{N} \sum_{t=k+1}^N y(t)y^*(t-k) \quad 0 \leq k \leq N-1$$

- Pour les deux estimateurs :

$$\hat{r}(-k) = \hat{r}^*(k), \quad k = 0, \dots, N-1$$

- **Commentaires :**

- Pour la plupart des signaux stationnaires, ACF décroît assez rapidement (valeurs assez petites pour des grand décalages  $k$ ).
- En pratique, on préfère l'estimateur biaisé (il possède de 'bonnes' propriétés pour la DSP)

# Relation entre Périodogramme et Corrélogramme

- Si on utilise la définition biaisé pour la covariance, alors il est possible de montrer que :

$$\begin{aligned}\hat{\phi}_p(\omega) &= \frac{1}{N} \left| \sum_{t=1}^N y(t) e^{-i\omega t} \right|^2 \\ &= \sum_{k=-(N-1)}^{N-1} \hat{r}(k) e^{-i\omega k} \\ &= \hat{\phi}_c(\omega)\end{aligned}$$

$$\boxed{\hat{\phi}_p(\omega) = \hat{\phi}_c(\omega)}$$



# Les performances

- Pour des petits  $N$ , biais important
- Pour des grands  $N$ , alors  $\hat{\phi}_p(\omega)$  **asymptotiquement non biaisé** :

$$E \left\{ \hat{\phi}_p(\omega) \right\} \rightarrow \phi(\omega) \text{ as } N \rightarrow \infty$$

- Mais la variance ne tends pas vers 0 (**l'estimateur n'est pas consistant**)
- Explications (intuitives) :
  - $\hat{r}(k) - r(k)$  peut être grande pour des valeurs importantes de  $|k|$
  - Même si les erreurs  $\{\hat{r}(k) - r(k)\}_{|k|=0}^{N-1}$  sont petites, l'erreur  $[\hat{\phi}_p(\omega) - \phi(\omega)]$  peut être grandes.

# Les extensions (quelques)

BE Matlab

- **Périodogramme simple (sans fenêtre de pondération)**

- Equivalent au corrélogramme (si estimateur ACF biaisé)
- Biais raisonnable mais variance très importante

$$y_j(t) = y((j-1)M + t) \quad t = 1, \dots, M$$

$$(j = 1, \dots, L \triangleq \lceil N/M \rceil)$$

- **Périodogramme modifié (avec fenêtre pondération)**

- Permet de réduire la variance au détriment d'une (légère) augmentation du biais.

$$\hat{\phi}_j(\omega) = \frac{1}{M} \left| \sum_{t=1}^M y_j(t) e^{-i\omega t} \right|^2$$

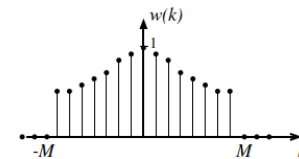
- **Bartlett, Welch estimations**

- Implémentation par moyennage du périodogramme (découpage par blocs sans/avec recouvrement) et sans/avec une fenêtre de pondération.
- Diminution de la variance (asymptotiquement non biaisé)

$$\hat{\phi}_B(\omega) = \frac{1}{L} \sum_{j=1}^L \hat{\phi}_j(\omega)$$

- **Blackman-Tukey estimation**

- Lissage / moyennage local du périodogramme par une fenêtre spectrale appropriée.
- Implémentation en tronquant et en pondérant l'ACF via une fenêtre adéquate

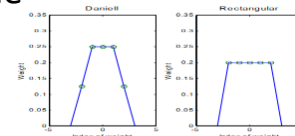


- **Daniell estimation**

- Approximativement l'approche de Blackman-Tukey avec une fenêtre spectrale de forme rectangulaire.
- Implémentation par filtrage/moyennage du périodogramme

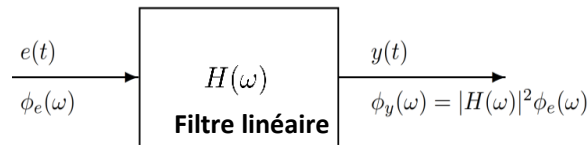
$$\hat{\phi}_{BT}(\omega) = \sum_{k=-(M-1)}^{M-1} w(k) \hat{r}(k) e^{-i\omega k}$$

$$W_n(j) = \begin{cases} \frac{1}{L} & \text{if } |j| < L/2, \\ 0 & \text{otherwise,} \end{cases}$$



# Conclusions

- Les méthodes non-paramétriques ne sont pas conçues pour inclure de l'information qui peut être disponible sur le processus pour l'estimation de son spectre.
- Que ce soit pour le périodogramme ou le corrélogramme, on tronque la fonction de covariance, d'où un fenêtrage, c'est-à-dire une perte de résolution et l'apparition de lobes secondaires (que l'on tente de filtrer par des fenêtres adéquates).
- Une solution est l'analyse spectrale paramétrique : un signal est modélisé comme la réponse d'un système à un autre signal de caractéristiques plus simples.

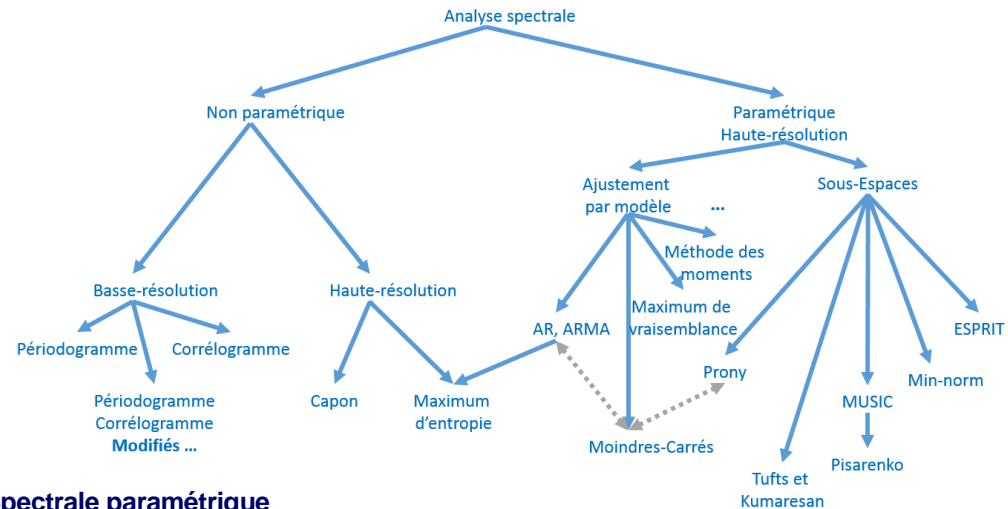


$$\phi_y(\omega) = |H(\omega)|^2 \phi_e(\omega)$$



Réponse fréquentielle du  
filtre

# Focus sur les Systèmes Linéaires Invariants (SLI)



# Quelques propriétés

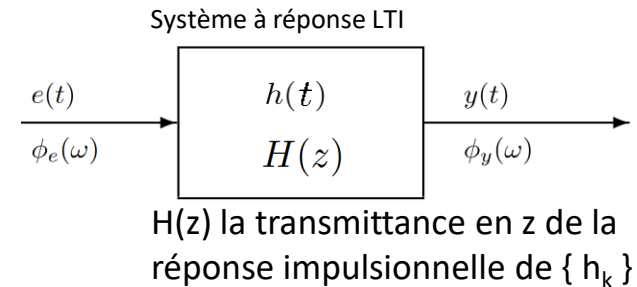
## • Modèle des Systèmes linéaires invariants

- Un système Linéaire Invariant dans le Temps (LTI) à temps discret est un modèle qui établit une relation de cause à effet entre l'entrée  $\{e(t)\}$  et la sortie  $\{y(t)\}$  via  $h(t) = \{h_k\}$ .
- $\{e(t)\}$  et  $\{y(t)\}$  sont relié par la convolution

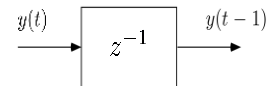
$$y(t) = h(t) * e(t) \quad \text{ou encore}$$

$$y(t) = H(z)e(t) = \sum_{k=-\infty}^{\infty} h_k e(t - k)$$

convolution !



$$Y(z) = H(z)E(z)$$

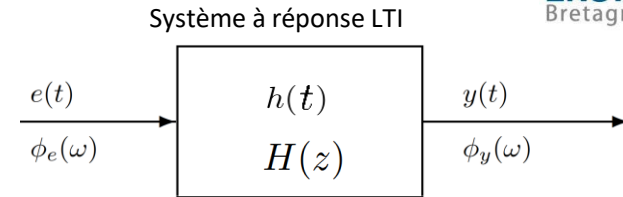


$$H(z) = \sum_{k=-\infty}^{\infty} h_k z^{-k} \quad \text{Fonction de transfert}$$

$z^{-1}$  désigne l'opérateur retard unité  
défini par :  $z^{-1}y(t) = y(t - 1)$

# Quelques propriétés

- Modèle des Systèmes linéaires invariants**



- Classiquement, les filtres LTI à temps discret peuvent s'exprimer à partir d'une équation récurrente de la forme :

$$y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = b_0 e(t) + b_1 e(t-1) + \dots + b_m e(t-m)$$

Les coefficients  $\{a_i\}_{i=1}^n$  sont associés à la partie récurrente du filtre

Les coefficients  $\{b_i\}_{i=0}^m$  sont associés à la partie non-récurrente du filtre

En prenant la transformée en  $z$  des deux membres :

$$z^{-1}y(t) = y(t-1)$$

$$Y(z) (1 + a_1 z^{-1} + \dots + a_n z^{-n}) = E(z) (b_0 + b_1 z^{-1} + \dots + b_m z^{-m})$$

$$\Rightarrow H(z) = \frac{Y(z)}{E(z)} = \frac{b_0 + b_1 z^{-1} + \dots + b_m z^{-m}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} = \frac{B(z)}{A(z)}$$

En fonction des paramètres  $\{a_i\}_{i=1}^n$  il est possible de définir deux catégories :

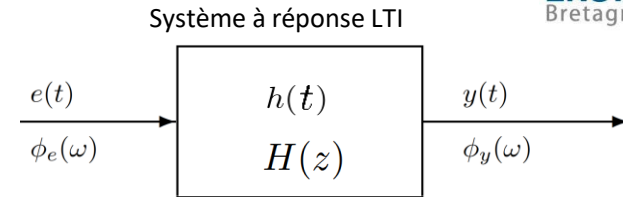
si  $A(z) = 1$  alors le filtre est dit Réponse Impulsionnelle Finie (FIR)

sinon si la fraction rationnelle  $H(z)$  est irréductible alors le filtre est dit Réponse Impulsionnelle Infinie (IIR)

La position des pôles de  $B(z)$  et des zéros de  $A(z)$  permettent de connaître la forme du gain de  $H(z)$

# Quelques propriétés

- Modèle des Systèmes linéaires invariants**



- Classiquement, l'étude d'un filtre passe par l'analyse dans le domaine fréquentielle de la fonction de transfert  $H(z)$ .

La Réponse fréquentielle de  $H(z)$  est obtenue en substituant  $z = e^{i\omega}$

$$H(z) = \sum_{k=-\infty}^{\infty} h_k z^{-k} \quad \Rightarrow \quad H(e^{i\omega}) = \sum_{k=-\infty}^{\infty} h_k e^{-i\omega k}$$

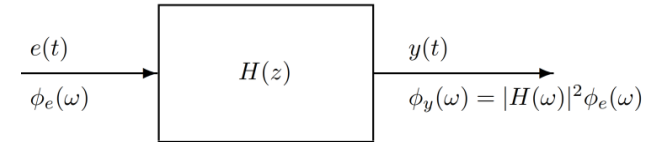
$$\boxed{\begin{aligned} H(e^{i\omega}) &= H(z)|_{z=e^{i\omega}} \\ &= H(\omega) \end{aligned}}$$

Ainsi si  $H(z)$  s'exprime sous la forme d'une fraction rationnelle

$$H(z) = \frac{B(z)}{A(z)} \quad \Rightarrow \quad H(\omega) = \frac{B(\omega)}{A(\omega)} = \frac{b_0 + b_1 e^{-i\omega} + \dots + b_m e^{-i\omega m}}{1 + a_1 e^{-i\omega} + \dots + a_n e^{-i\omega n}}$$

# Quelques propriétés

- Relation entre DSP et système linéaire



On considère

$$y(t) = \sum_{k=-\infty}^{\infty} h_k e(t-k)$$

Expression de :  $r_y(k) = E\{y(t)y^*(t-k)\}$

$$\begin{aligned} r_y(k) &= \sum_{p=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_p h_m^* E\{e(t-p)e^*(t-m-k)\} \\ &= \sum_{p=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_p h_m^* r_e(m+k-p) \end{aligned}$$

Puis en prenant la TFTD :

$$\begin{aligned} \phi_y(\omega) &= \sum_{k=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_p h_m^* r_e(m+k-p) e^{-i\omega(k+m-p)} e^{i\omega m} e^{-i\omega p} \\ &= \left[ \sum_{p=-\infty}^{\infty} h_p e^{-i\omega p} \right] \left[ \sum_{m=-\infty}^{\infty} h_m^* e^{i\omega m} \right] \left[ \sum_{\tau=-\infty}^{\infty} r_e(\tau) e^{-i\omega \tau} \right] \\ &= |H(\omega)|^2 \phi_e(\omega) \end{aligned}$$

$\Rightarrow$

$$\phi_y(\omega) = |H(\omega)|^2 \phi_e(\omega)$$

Ou encore

$\Rightarrow$

$$\phi_y(z) = H(z) H^*(1/z^*) \phi_e(z)$$



# Le problème d'estimation spectrale

- **Le problème**

A partir d'un signal discret de  $N$  échantillons (une réalisation avec souvent  $N$  petit)

$$\{y(1), \dots, y(N)\}$$

Comment trouver une estimation de la DSP via un modèle paramétrique :

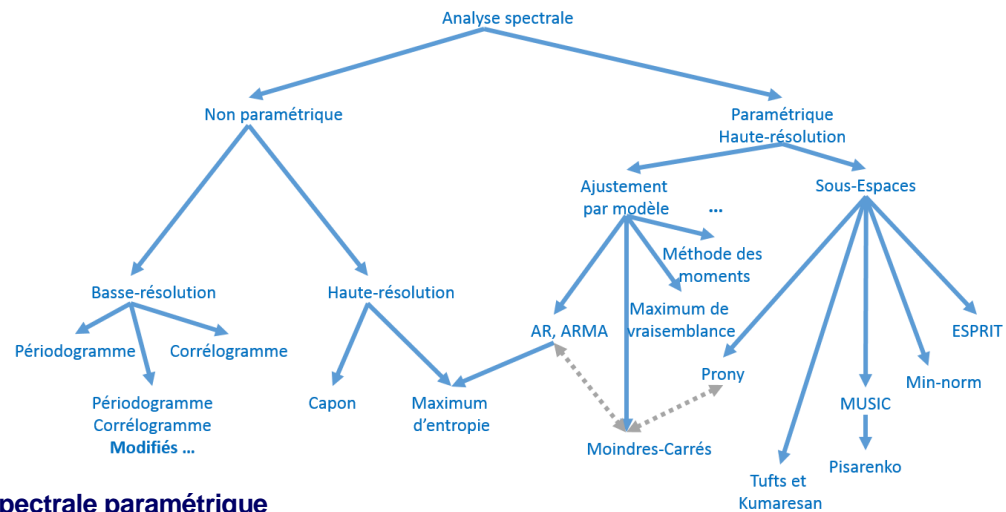
$$\{\hat{\phi}(\omega), \omega \in [-\pi, \pi]\}$$

- **Solutions**

Via deux approches différentes :

- Les approches non paramétriques qui découlent des définitions de la DSP
- Les approches paramétriques qui se base sur une modélisation (paramétrée) de la DSP

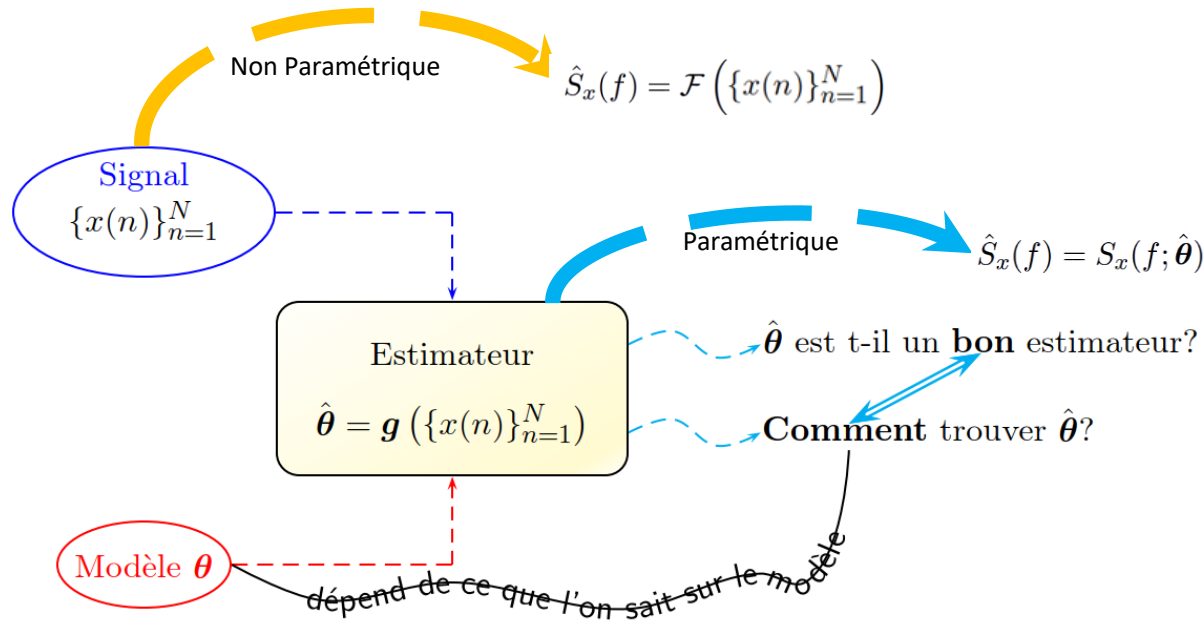
# Modèles Rationnels & Factorisation spectrale



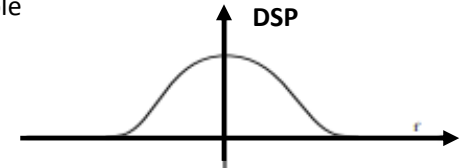
Estimation Spectrale paramétrique

# Introduction

## • Modélisation paramétrique



Exemple

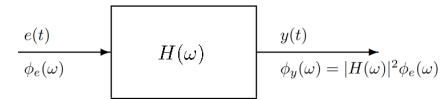


$$P(f) = \frac{r(0)}{\sqrt{2\pi}\sigma_f} e^{-\frac{1}{2}\left(\frac{f}{\sigma_f}\right)^2}, |f| \leq \frac{1}{2}$$

- La DSP est décrit par 2 inconnus :  $\sigma_f$   $r(0)$

- Une fois que nous connaissons les deux, nous connaissons la DSP
- Les méthodes paramétriques tentent d'estimer les deux paramètres.

# Modèles rationnels



On souhaite décrire une DSP par un modèle rationnel linéaire de la forme

$$\phi(\omega) = \frac{\sum_{|k| \leq m} \gamma_k e^{-i\omega k}}{\sum_{|k| \leq n} \rho_k e^{-i\omega k}} \quad (\text{Théorème de Weierstrass})$$

$\phi(\omega)$  est une fonction rationnel en  $e^{-i\omega}$

Plus précisément, puisque  $\phi(\omega) \geq 0$ :

$$\phi(\omega) = \left| \frac{B(\omega)}{A(\omega)} \right|^2 \sigma^2$$

$$B(\omega) = 1 + b_1 e^{-i\omega} + \dots + b_m e^{-im\omega}$$

$$A(\omega) = 1 + a_1 e^{-i\omega} + \dots + a_n e^{-in\omega}$$

$\sigma^2$  un scalaire positif

# Modèles rationnels

## Théorème de factorisation spectrale

$$\phi(\omega) = \left| \frac{B(\omega)}{A(\omega)} \right|^2 \sigma^2$$

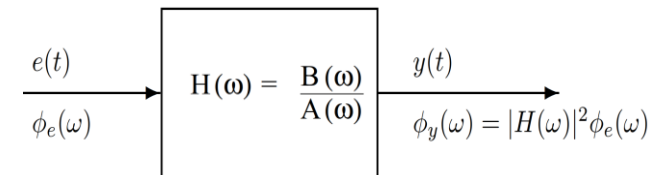
$$B(\omega) = 1 + b_1 e^{-i\omega} + \dots + b_m e^{-im\omega}$$

$$A(\omega) = 1 + a_1 e^{-i\omega} + \dots + a_n e^{-in\omega}$$

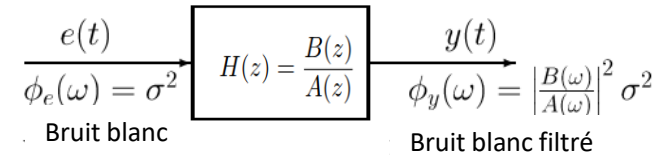
$\sigma^2$  un scalaire positif

Une DSP rationnelle arbitraire peut être associée à un signal obtenu en filtrant un bruit blanc, centré, de puissance  $\sigma^2$  à travers un filtre rationnel dont la fonction de transfert est  $H(\omega)$  :

$$H(\omega) = \frac{B(\omega)}{A(\omega)}$$



# Modèles rationnels



Transformé en  $z$  de  $H$  (en posant  $z = e^{i\omega}$ ) :

$$H(z) = \frac{B(z)}{A(z)}$$

$$A(\omega) = A(z)|_{z=e^{i\omega}}$$

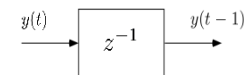
$$B(\omega) = B(z)|_{z=e^{i\omega}}$$

$$B(z) = 1 + b_1 z^{-1} + \dots + b_m z^{-m}$$

$$A(z) = 1 + a_1 z^{-1} + \dots + a_n z^{-n}$$

En temps, cela revient à écrire la relation de convolution entre  $\{e(t)\}$  et  $\{y(t)\}$  :

$$y(t) = H(z)e(t) = \frac{B(z)}{A(z)}e(t)$$



$z^{-1}$  opérateur retard unité

$$z^{-1}y(t) = y(t-1)$$

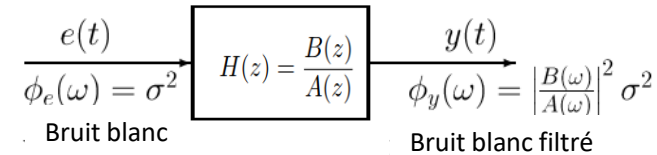
Ou encore

$$A(z)y(t) = B(z)e(t)$$

$e(t)$  bruit blanc de variance  $\sigma^2$

$y(t)$  sortie du filtre

# Modèles rationnels



**Equation de récurrence linéaire d'ordre fini, à coefficients constants :**

$$y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{j=0}^m b_j e(t-j), \quad (b_0 = 1)$$

$$A(z)y(t) = B(z)e(t)$$

**Modèles de signal :**

- Auto Regressive Moving Average ARMA(n,m)

$$\phi(\omega) = \sigma^2 \left| \frac{B(\omega)}{A(\omega)} \right|^2 \quad y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{j=0}^m b_j e(t-j), \quad (b_0 = 1)$$

- **Auto Regressive AR(n)**

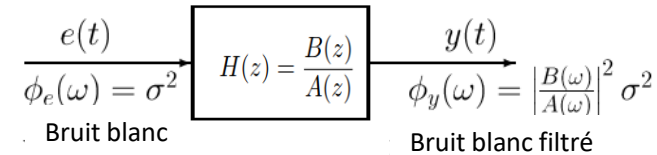
$$\phi(\omega) = \sigma^2 \left| \frac{1}{A(\omega)} \right|^2 \quad y(t) + \sum_{i=1}^n a_i y(t-i) = e(t)$$

- Moving Average MA(m)

$$\phi(\omega) = \sigma^2 |B(\omega)|^2 \quad y(t) = \sum_{j=0}^m b_j e(t-j), \quad (b_0 = 1)$$

# Factorisation Spectrale

$$\begin{aligned}\phi_y(z) &= |H(z)|^2 \phi_e(z) \\ &= H(z)H^*(1/z^*)\phi_e(z) \\ &= \frac{B(z)B^*(1/z^*)}{A(z)A^*(1/z^*)} \sigma^2\end{aligned}$$



$$A(\omega) = A(z)|_{z=e^{i\omega}}$$

$$B(\omega) = B(z)|_{z=e^{i\omega}}$$

$$\begin{aligned}B(z) &= 1 + b_1 z^{-1} + \dots + b_m z^{-m} \\ A(z) &= 1 + a_1 z^{-1} + \dots + a_n z^{-n}\end{aligned}$$

## • Quelques propriétés (pont de vue filtre) :

- Si  $z_i = r e^{i\theta}$  est un zéros / pôles alors  $(1/z_i^*) = (1/r) e^{i\theta}$  aussi.
- Pour que le processus soit stationnaire,  $H(z) = \frac{B(z)}{A(z)}$  doit être stable et causal
  - Si les racines de  $A(z)$  (pôles de  $\frac{1}{A(z)}$ ) sont à l'intérieure du cercle unité alors  $H(z)$  stable
  - Si les racines de  $B(z)$  (zéros de  $H(z)$ ) sont à l'intérieure du cercle unité alors  $H(z)$  est à phase minimale.
- Si  $\{a_i\}_{i=1}^n$  et  $\{b_i\}_{i=1}^m$  sont réels alors  $\phi_y(z) = \sigma^2 \frac{B(z)B(\frac{1}{z})}{A(z)A(\frac{1}{z})}$  (les racines / pôles sont complexe conjugués).



# Factorisation Spectrale

$$\phi(\omega) = \left| \frac{B(\omega)}{A(\omega)} \right|^2 \sigma^2$$

$$\begin{aligned} B(\omega) &= 1 + b_1 e^{-i\omega} + \dots + b_m e^{-im\omega} \\ A(\omega) &= 1 + a_1 e^{-i\omega} + \dots + a_n e^{-in\omega} \end{aligned}$$

## • Influence de A(z) et B(z)

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^m b_k z^{-k}}{\sum_{k=0}^n a_k z^{-k}} \quad (a_0 = 1, b_0 = 1)$$

La DSP dépend entièrement des racines de A(z) et B(z) :

**Les pôles de H(z) :**

$$A(z) = \prod_{k=1}^n (1 - p_k z^{-1}) = \prod_{k=1}^n (1 - \rho_k e^{-i\omega_k} z^{-1})$$

Les pôles  $p_k$  correspondent à des **pics** dans la DSP, situés à  $\frac{\omega_k}{2\pi}$ .

**Pics** d'autant plus fort que  $\rho_k$  est proche de 1, c'est-à-dire que  $p_k$  est proche du cercle unitaire.

**Les zéros de H(z) :**

$$B(z) = \prod_{k=1}^m (1 - z_k z^{-1}) = \prod_{k=1}^m (1 - \varphi_k e^{-i\psi_k} z^{-1})$$

Les zéros  $z_k$  correspondent à des **vallées** dans la DSP, situés à  $\frac{\psi_k}{2\pi}$ .

**Vallées** d'autant plus fort que  $\varphi_k$  est proche de 1, c'est-à-dire que  $z_k$  est proche du cercle unitaire.

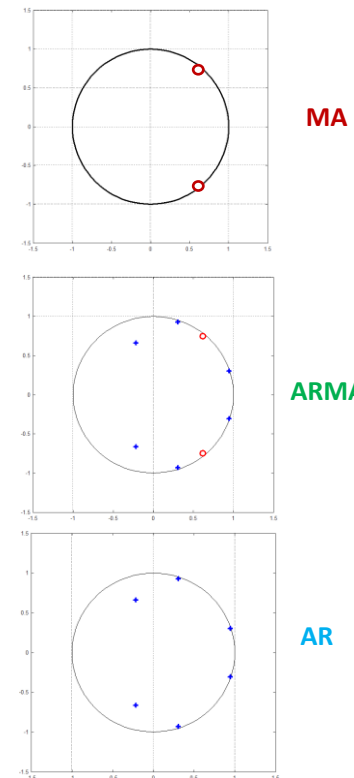
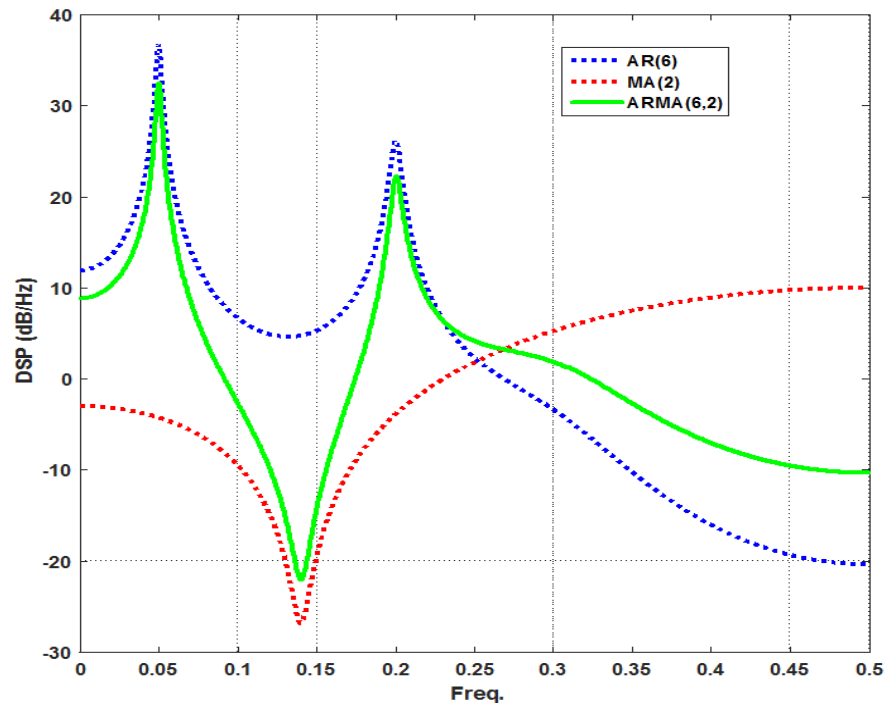
⇒ Ainsi, un processus ARMA(n, m) est capable d'approximer très précisément une DSP en fonction de n et m.

# Factorisation spectrale

- Exemple de DSP (théorique) : AR(n), MA(m), ARMA(n,m)

Zeroes :  $0.97 e^{\pm i2\pi 0.14}$

Poles :  $0.99 e^{\pm i2\pi 0.05}$   $0.98 e^{\pm i2\pi 0.2}$   $0.7 e^{\pm i2\pi 0.3}$



# Relations entre les modèles

## • Théorème de décomposition de Wold (conséquence)

- **Tout modèle MA(m) ou ARMA(n, m) est équivalent à un modèle AR( $\infty$ )**
- Tout modèle AR(n) ou ARMA(n, m) est équivalent à un modèle MA( $\infty$ )

$$\text{ARMA}(n, m) = \frac{B(z)}{A(z)} = \frac{1}{C(z)} = \text{AR}(\infty).$$

**Démonstration :**

$$\begin{aligned} B(z) &= 1 + b_1 z^{-1} + \dots + b_m z^{-m} \\ A(z) &= 1 + a_1 z^{-1} + \dots + a_n z^{-n} \end{aligned}$$

De  $\{a_i\}_{i=1}^n$  et  $\{b_i\}_{i=1}^m$ , il est possible d'exprimer  $\{c_i\}_{i=1}^\infty$

Puisque

$$\begin{aligned} \frac{B(z)}{A(z)} = \frac{1}{C(z)} &\Rightarrow C(z)B(z) = A(z) \\ \Rightarrow [1 + c_1 z^{-1} + \dots][1 + b_1 z^{-1} + \dots + b_m z^{-m}] &= [1 + a_1 z^{-1} + \dots + a_n z^{-n}] \end{aligned}$$

Puis par identification, il est possible d'exprimer les  $\{c_i\}_{i=1}^\infty$ .

N.B. : Il est possible d'avoir une écriture de type matricielle de la forme :

$$\begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ c_1 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ c_n & \ddots & \ddots & \ddots & \\ c_{n+1} & \ddots & \ddots & \ddots & 1 \\ \vdots & \ddots & \ddots & \ddots & c_1 \\ \vdots & & & & c_n \\ & & & & \vdots \end{bmatrix} \begin{bmatrix} 1 \\ b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_n \\ 0 \\ \vdots \end{bmatrix}$$

# Relation avec la covariance (ACF)

$$\phi(\omega) = \left| \frac{B(\omega)}{A(\omega)} \right|^2 \sigma^2$$

$$\begin{aligned} B(\omega) &= 1 + b_1 e^{-i\omega} + \dots + b_m e^{-im\omega} \\ A(\omega) &= 1 + a_1 e^{-i\omega} + \dots + a_n e^{-in\omega} \end{aligned}$$

La spectre d'un modèle ARMA(n, m) dépend uniquement des paramètres :  $\{a_i\}_{i=1}^n \quad \{b_i\}_{i=1}^m \quad \sigma^2$

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k}$$

Par conséquent, la fonction de covariance (ACF) dépend également que de ces paramètres

On a en ainsi pour un modèle ARMA(n, m) :

$$y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{j=0}^m b_j e(t-j), \quad (b_0 = 1)$$

$$y(t) = H(z)e(t) = \sum_{k=0}^{\infty} h_k e(t-k).$$

$$r(k) = E \{y(t)y^*(t-k)\}$$

$$r(k) = \begin{cases} r^*(-k) & k < 0 \\ -\sum_{i=1}^n a_i r(k-i) + \sigma^2 \sum_{j=0}^m b_j h_{j-k}^* & k \in [0, m] \\ -\sum_{i=1}^n a_i r(k-i) & k > m \end{cases}$$

⇒ relations de Yule-Walker pour un processus ARMA(n, m).

# Relation avec la covariance (ACF)

$$r(k) = \begin{cases} r^*(-k) & k < 0 \\ -\sum_{i=1}^n a_i r(k-i) + \sigma^2 \sum_{j=0}^m b_j h_{j-k}^* & k \in [0, m] \\ -\sum_{i=1}^n a_i r(k-i) & k > m \end{cases}$$

relations de Yule-Walker  
d'un processus ARMA(n, m)

La majorité des méthodes d'estimation des paramètres ARMA, AR, MA  
est basée sur les relations de Yule-Walker.

Dans le cadre d'un processus AR(n), il est possible d'avoir une expression simple de la dernière ligne de la forme, ( $k>0$  et  $m=0$ ) :

**R**

$$\begin{bmatrix} r(1) \\ \vdots \\ r(n) \end{bmatrix} = - \begin{bmatrix} r(0) & \cdots & r(-n+1) \\ \vdots & \ddots & \vdots \\ r(n-1) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

$$\mathbf{R}\mathbf{a} = -\mathbf{r}$$

# Relation avec la covariance (ACF)

$$r(k) = \begin{cases} r^*(-k) & k < 0 \\ -\sum_{i=1}^n a_i r(k-i) + \sigma^2 \sum_{j=0}^m b_j h_{j-k}^* & k \in [0, m] \\ -\sum_{i=1}^n a_i r(k-i) & k > m \end{cases}$$

relations de Yule-Walker  
d'un processus ARMA(n, m)

La majorité des méthodes d'estimation des paramètres ARMA, AR, MA  
est basée sur les relations de Yule-Walker.

**Dans le cadre d'un processus AR(n)**, il est possible d'avoir une expression simple de la dernière ligne de la forme, ( $k>0$  et  $m=0$ ) :

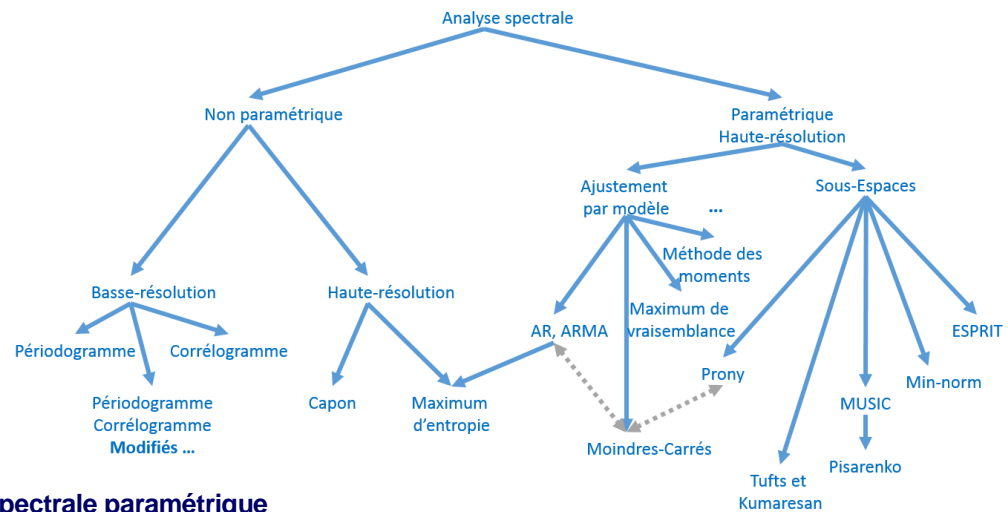
Ou encore :

**Equation normale**

**Yule Walker**

$$\begin{bmatrix} r(0) & r(-1) & \dots & r(-n) \\ r(1) & r(0) & & \vdots \\ \vdots & & \ddots & r(-1) \\ r(n) & \dots & & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

# Modèles AR & Analyse spectrale



# Estimation du spectre avec un modèle AR

- Un processus AR(n) est la sortie d'un filtre RII (Réponse impulsionnelle infinie) dont l'entrée est un bruit blanc de variance  $\sigma^2$

$$y(t) + \sum_{i=1}^n a_i y(t-i) = e(t) \quad (b_0 = 1)$$

- Le spectre d'un processus AR(n) est

$$\phi(\omega) = \frac{\sigma^2}{|1 + \sum_{i=1}^n a_i \exp(-j\omega i)|^2}$$

- Ainsi, *si les paramètres peuvent être estimés à partir des données disponibles*, alors l'estimation du spectre est

$$\hat{\phi}_{\text{AR}}(\omega) = \frac{\hat{\sigma}^2}{|1 + \sum_{i=1}^n \hat{a}_i \exp(-j\omega i)|^2}$$

- La précision du spectre dépendra de la précision de l'estimation des paramètres du modèle ainsi que du choix du modèle lui-même.
- Pour estimer les paramètres d'un processus AR, on peut estimer les fonctions de corrélations (ACF) et à partir de ces estimées, par utilisation des équations de Yule-Walker, il est alors possible d'en déduire les paramètres.



# Processus AR

 $\Rightarrow$ 

$$\begin{bmatrix} r(0) & r(-1) & \dots & r(-n) \\ r(1) & r(0) & & \vdots \\ \vdots & & \ddots & r(-1) \\ r(n) & \dots & & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Equations de Yule-Walker ou équations normal

Si les  $\{r(k)\}_{k=0}^n$  sont estimés en utilisant l'estimateur biaisé ACF

$$\Rightarrow \begin{bmatrix} r(1) \\ \vdots \\ r(n) \end{bmatrix} + \begin{bmatrix} r(0) & \dots & r(-n+1) \\ \vdots & \ddots & \vdots \\ r(n-1) & \dots & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow$$

La solution est

$$\theta_n = -R_n^{-1} r_n$$

L'estimation de  $\sigma^2$  est faite en écrivant  $r(0) + \sum_{i=1}^n a_i r(-i) = \sigma^2$

$$r_n + R_n \theta_n = 0$$

$$\theta_n = [a_1, \dots, a_n]^T$$

$$R_{n+1} \begin{bmatrix} 1 \\ \theta_n \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \end{bmatrix}$$

# Processus AR

$$\hat{\theta}_n = -\hat{R}_n^{-1} \hat{r}_n$$

## Remarques

- Si N échantillons alors  $\{r(k)\}_{k=0}^n$  n'est pas disponible mais  $\{\hat{r}(k)\}_{k=0}^n$  peut être utilisé pour obtenir  $\sigma^2$  et  $\{\hat{a}_i\}_{i=0}^n$   
 $\Rightarrow$  C'est la méthode de Yule-Walker.
- La matrice  $R_{n+1}$  a une structure très particulière. Elle est notamment hermitienne et de Toëplitz.
- L'estimation des paramètres de l'AR est un sujet bien établi dans la littérature et des algorithmes existent pour résoudre de manière efficace l'équation :
  - Algorithme récursif de Levinson-Durbin (1947, 1960)
  - Algorithme récursif de Delsarte-Genin (1986) dans le cas réels
  - Algorithme récursif de Burg (1975)
  - **AlgorithmeS des moindres carrés (Least-Squares method)**
- Les modèles AR sont les plus fréquemment utilisés dans la pratique.

# Processus AR

## • Algorithme des moindres carrés (Least-Squares method)

$$y = - \begin{bmatrix} y(n+1) \\ y(n+2) \\ \vdots \\ y(N) \end{bmatrix}$$

$$M = \begin{bmatrix} y(n) & y(n-1) & \dots & y(1) \\ y(n+1) & y(n) & \dots & y(2) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & \dots & y(N-n) \end{bmatrix}$$

$$\hat{\theta} = -(M^T M)^{-1} M^T y$$

Algorithme des moindres carrés (méthode des covariances) :

### On note

$n$  : Ordre du processus AR

$N$  : Nombre d'échantillons de  $y(t)$

### Etape 1 :

Construction de  $y$  et de la matrice  $M$

$y1 = y(n+1:N)$

$M = \text{toeplitz}(y(n:N-1), y(n:-1:1).')$  (commande Matlab)

### Etape 2 :

Estimation  $\hat{a} = \{a_i\}_{i=1}^n$  par moindres carrés

$$\hat{a} = -(M^T M)^{-1} M^T y$$

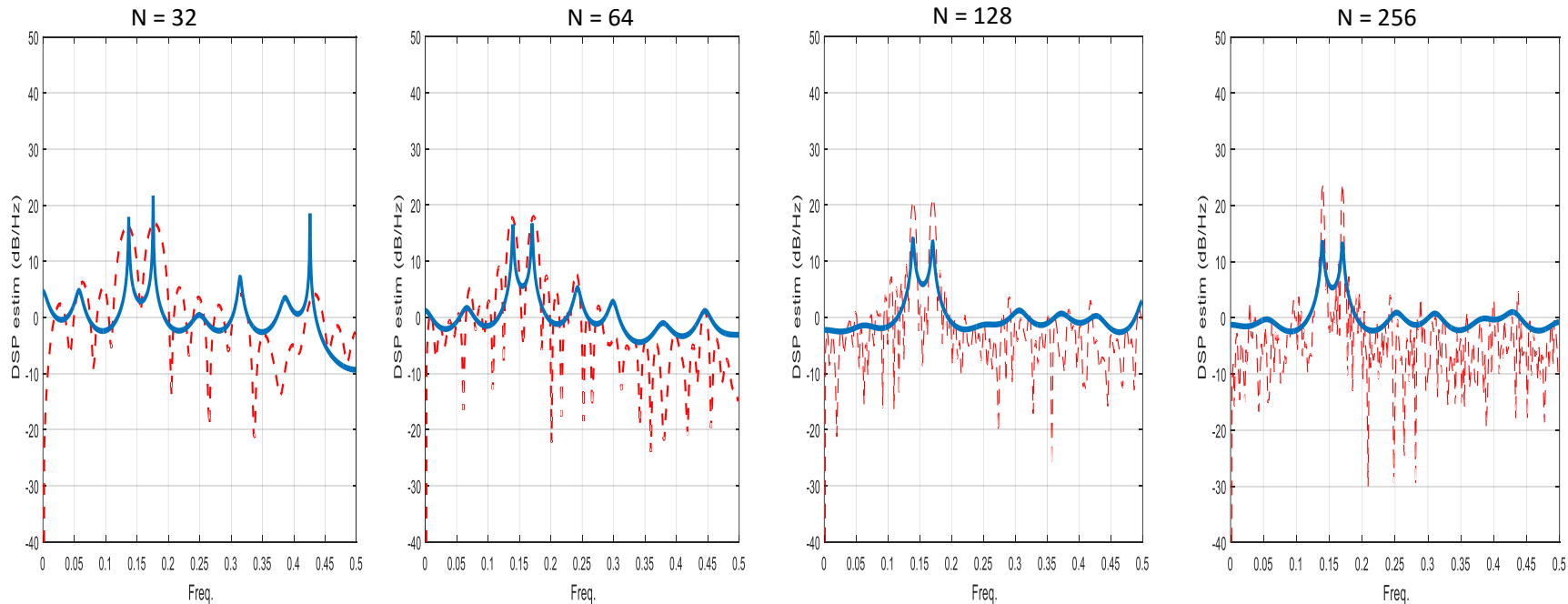
Estimation du bruit

$$\hat{\sigma} = \frac{1}{N-n} \sum r^2 \quad \text{avec} \quad r = M\hat{a} + y1 \quad (r: \text{le résidu})$$

# Processus AR

- **Le pouvoir « haute résolution » de l'estimateur AR.**

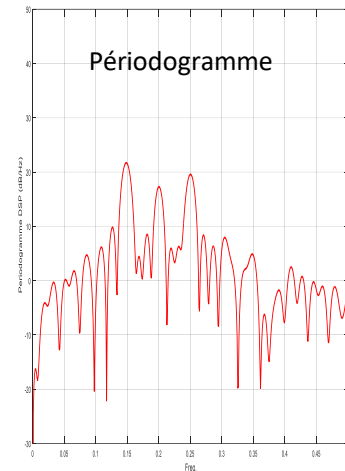
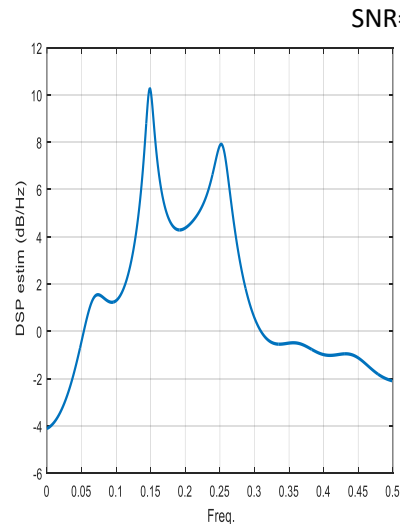
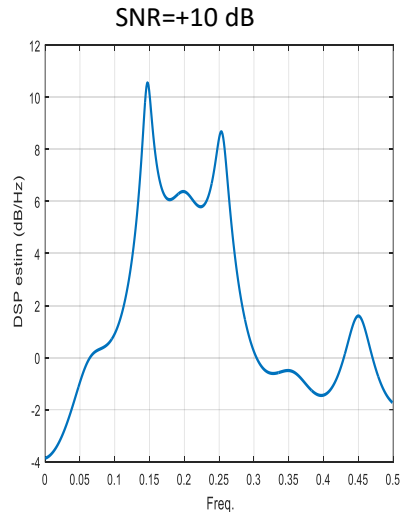
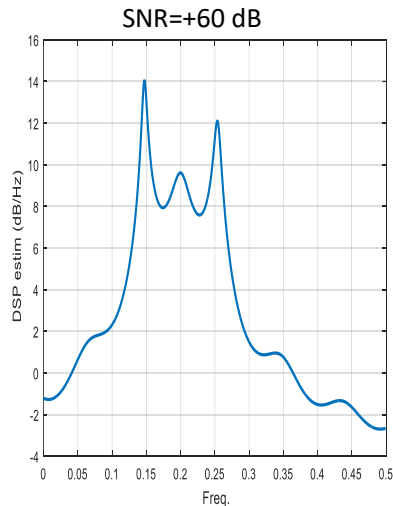
Dans cet exemple, estimation de 3 sinusoïdes (Amplitudes : 1,5 0,9 1,1 & Fréquences : 0,15 0,2 0,25) noyées dans un bruit blanc gaussien (SNR=10 dB). Modélisation par un processus AR(16) et périodogramme. N: Nombre d'échantillons



# Processus AR

- **Le pouvoir « haute résolution » de l'estimateur AR.**

Dans cet exemple, estimation de 3 sinusoïdes (Amplitudes : 1,5 0,9 1,1 & Fréquences : 0,15 0,2 0,25) noyées dans un bruit blanc gaussien. Nombre de points :  $N=55$ . Modélisation par un processus AR(12).



Notez ce qui se passe dans le dernier cas. L'estimateur nous donne une solution dégénérée contenant seulement 2 sinusoïdes. Les sinusoïdes (sources) sont trop « proches » pour pouvoir être « distinguée » à ce niveau de SNR. Il faudrait utiliser un modèle AR d'ordre supérieure ou alors un ARMA.

# Processus AR

## Conclusion : un modèle rationnel linéaire universel

- Plus haute résolution que les méthodes classiques : la possibilité d'extrapoler la fonction de corrélation, et donc d'éliminer le fenêtrage élimine les lobes secondaires.
  - Algorithmes rapides.
  - Bonne performance pour des signaux de courte durée.
- 
- Difficulté à choisir l'ordre du modèle.
  - Résolution décroissante avec le SNR.
  - Risque de doublement de raies dans certains cas.
  - Contrairement au périodogramme, pour un signal sinusoïdal, l'amplitude du pic AR n'est pas directement proportionnelle à la puissance de la sinusoïde.
  - Dans le cas d'une somme de  $n$  sinusoïdes bruitées, une modélisation AR(L) (avec  $L > 2n$ ) associée à une recherche des  $n$  pôles (ou maxima) prépondérants permet d'obtenir une bonne estimation des fréquences lorsque le SNR est "bon". Néanmoins, ces estimateurs restent biaisés et on peut même montrer que le biais est de l'ordre de  $O(1/L^3)$

- Un processus MA(m) est la sortie d'un filtre RIF (Réponse impulsionnelle finie) dont l'entrée est un bruit blanc de variance  $\sigma^2$

$$y(t) = \sum_{k=0}^m b_k e(t-k), \quad (b_0 = 1)$$

- Le spectre d'un processus MA(m) est

$$\phi(\omega) = \sigma^2 \left| \sum_{k=0}^m b_k \exp(-j\omega k) \right|^2$$

- Si les paramètres peuvent être estimés, alors l'estimation du spectre est

$$\hat{\phi}_{MA}(\omega) = \hat{\sigma}^2 \left| \sum_{k=0}^m \hat{b}_k \exp(-j\omega k) \right|^2$$

- La précision du spectre dépendra bien évidemment de la précision de l'estimation des paramètres du modèle ainsi que du choix du modèle lui-même.

- Pour estimer les paramètres d'un processus MA, il est possible d'utiliser :

- Méthode de Durbin : utiliser le fait qu'un MA peut être modéliser par un AR(L=∞ )  $\sum_{k=0}^m b_k z^{-k} = \frac{1}{A_\infty(z)} \approx \frac{1}{\sum_{k=0}^L a_k z^{-k}}$  [en réalité L >> m] puis utilisation des relations entre les modèles.

- Utiliser le fait que le spectre peut aussi s'écrire selon  $\phi(\omega) = \sum_{k=-m}^m r(k) e^{-i\omega k}$  et avec  $r(k) = 0$  for  $|k| > m$  (équivalent à l'estimateur de Blackman-Tukey pour un processus MA)

# Retour sur l'estimation du spectre avec un modèle ARMA

- Un processus ARMA(n, m) est la sortie d'un filtre dont l'entrée est un bruit blanc de variance  $\sigma^2$

$$H(z) = \frac{\sum_{k=0}^m b_k z^{-k}}{1 + \sum_{k=1}^n a_k z^{-k}} \quad y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{j=0}^m b_j e(t-j), \quad (b_0 = 1)$$

- Le spectre d'un processus ARMA(n, m) est

$$\phi(\omega) = \sigma^2 \frac{|\sum_{k=0}^m b_k \exp(-j\omega k)|^2}{|1 + \sum_{k=1}^n a_k \exp(-j\omega k)|^2}$$

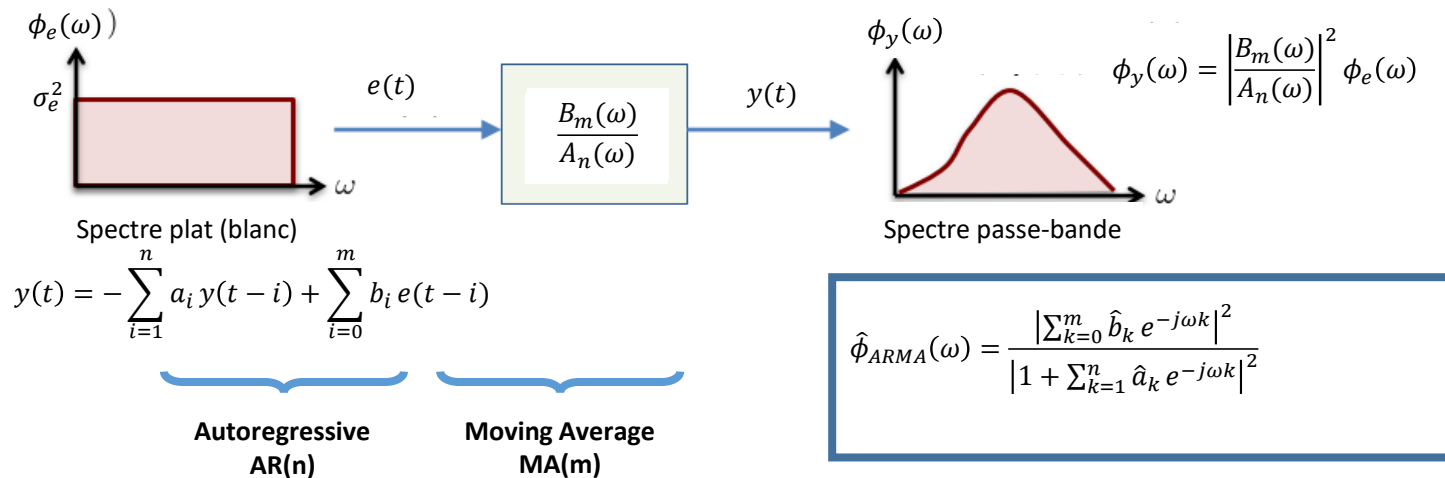
- L'estimation des paramètres ARMA se fait classiquement en deux temps : tout d'abord, estimation des paramètres AR et ensuite estimation des paramètres MA. Cette solution sous-optimale conduit à une réduction de la complexité calculatoire (par rapport à la solution optimale qui serait d'estimer conjointement les paramètres AR et MA) (méthode de Yule-Walker modifiées).
- Le spectre d'un processus ARMA peut être estimée, une fois que les paramètres AR et MA sont estimés

$$\hat{\phi}(\omega) = \hat{\sigma}^2 \frac{|\sum_{k=0}^m \hat{b}_k \exp(-j\omega k)|^2}{|1 + \sum_{k=1}^n \hat{a}_k \exp(-j\omega k)|^2}$$



# Estimation spectrale par modèles rationnels

- Un ARMA (n, m) permet d'approcher très précisément la DSP d'une grande classe de signaux. La partie AR modélise les pics du spectre tandis que la partie MA modélise les vallées.
- Les paramètres du modèle sont généralement estimés en résolvant les équations de Yule-Walker (qui impliquent la fonction de corrélation). Ces équations sont linéaires par rapport aux paramètres AR, non linéaires par rapport aux paramètres MA.
- Les informations sur le contenu spectral peuvent être extraites de la DSP estimé par modèle ARMA (rationnel) ou encore par de l'examen des pôles et des **zéros** du modèle ARMA.
- Pour un modèle AR(n), la résolution des équations de Yule-Walker équivaut à la minimisation de l'erreur de prédiction linéaire.
- Les modèles AR et ARMA permettent l'estimation de la fréquence de signaux exponentiels complexes, le modèle ARMA offrant une meilleure résolution.



# Pour finir : retour sur la notion de modèles

L'idée de modélisation est de proposer, classiquement, une équation dynamique d'évolution permettant de rendre compte de l'évolution temporelle du signal à travers un modèle compacte (principe de parcimonie) et appropriée pouvant être :

## Modèles linéaires stationnaires :

⇒ AR, MA, ARMA

## Modèles linéaires non stationnaires (spectre évolutif) :

⇒ ARIMA, SARIMA, ARMA évolutif (Box et Jenkins 1976)

## Modèles exponentiels (complexes) :

⇒ approche déterministe (bruité ou pas) (prony harmonique)

⇒ approche signaux aléatoires stationnaires

## Modèles non-linéaires :

⇒ extension non-linéaire ARMA : . Granger et Anderson (1978) Tong (1978).

⇒ Modélisation ARCH /GARCH Engle (1982)

⇒ Relation entre les paramètres du modèle et la DSP ne sont pas toujours, à ce jour, déterminé. Priestley (1991) ou Tong (1990).

Fin !

