

analyse spectrale des séries temporelles



Jacques Beau

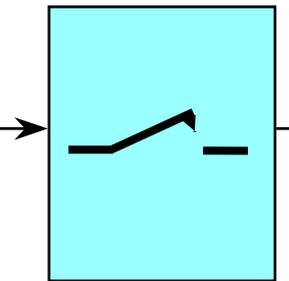
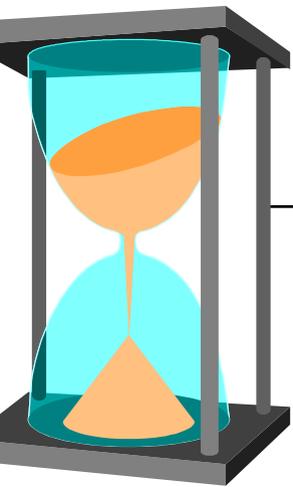
Rythmes biologiques et cancer, Inserm U 776

Hôpital Paul Brousse. 948000 Villejuif

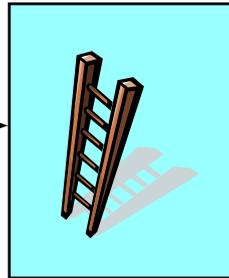
déroulement de la présentation

- ➡ **saisie des données numériques : échantillonnage et fenêtrage**
- ➡ **composantes spectrales**
- ➡ **relation entre domaines spectral et temporel**
- ➡ **transformation par FFT (*Fast Fourier Transform*)**
 - **définitions générales**
 - **ajout de zéros pour accroître la discrimination fréquentielle**
 - **fenêtres de pondération**
 - **les erreurs de mesure**
- ➡ **analyse spectrale par fenêtre glissante**
- ➡ **conclusion**

saisie des données chronobiologiques



échantillonnage



quantification



résultats
→ &
statistique

l'échantillonnage fournit une série numérique qui implique une double discrétisation :

- dans le temps : l'échantillonnage, caractérisé par sa période d'échantillonnage : $T_s = 1/F_s$
- dans l'espace : la quantification, caractérisée par le nombre de niveaux distincts de mesure ; avec 8 bits on obtient une échelle de 256 niveaux entre 0 et 255 en UA (non abordé ici).

en outre, la durée de la mesure étant finie, une fenêtre d'observation T_w est de fait appliquée au phénomène étudié.

Acquisition des données : échantillonnage - fenêtre

théorème de Shannon

$$F_e \geq 2F_{\max}$$

exemples

analyse d'un phénomène circadien (# 24 h)

la durée T_e séparant 2 échantillons doit être $T_e < 12$ h

analyse du rythme cardiaque (max 300 bat/min)

la durée T_e séparant 2 échantillons doit être $T_e < 0.2$ "

remarques

cette limite suppose un phénomène **parfaitement sinusoïdal**, ce qui n'est jamais vérifié; en pratique un facteur 2 ou 3 doit être appliqué pour tenir compte d'éventuelles périodes harmoniques.

la résolution dans la mesure d'une période est liée à la durée de la mesure

$$\Delta f_m \approx \frac{1}{T_w}$$

$$T_w 3 = 24 * 3 \Rightarrow T_{\max} = 35.99 \text{ et } T_{\min} = 18.00$$

$$T_w 10 = 24 * 10 \Rightarrow T_{\max} = 26.66 \text{ et } T_{\min} = 21.81$$

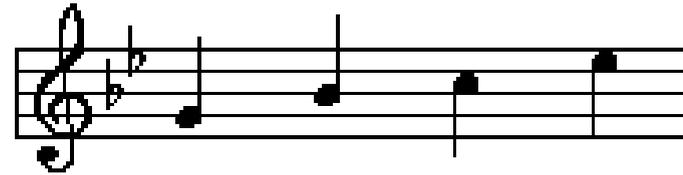
la précision de la mesure peut cependant être bien meilleure

l'analyse spectrale - composantes

relation entre les domaines temporel et fréquentiel

série de Fourier et transformée de Fourier

- la série : phénomène périodique
 - les composantes sont des raies harmoniques
- la transformée : phénomène limité dans le temps
 - analyse globale
 - analyse par fenêtre glissante
 - transformation de Fourier rapide (FFT)



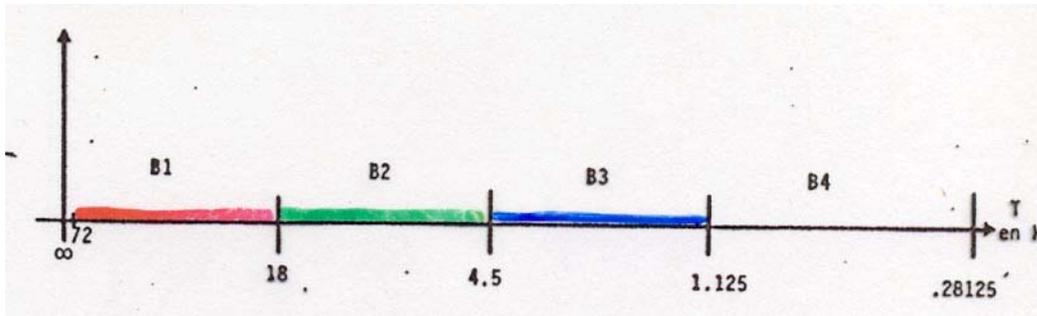
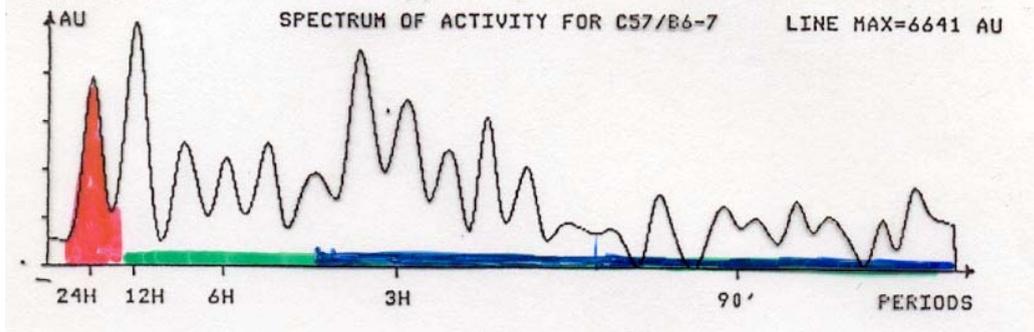
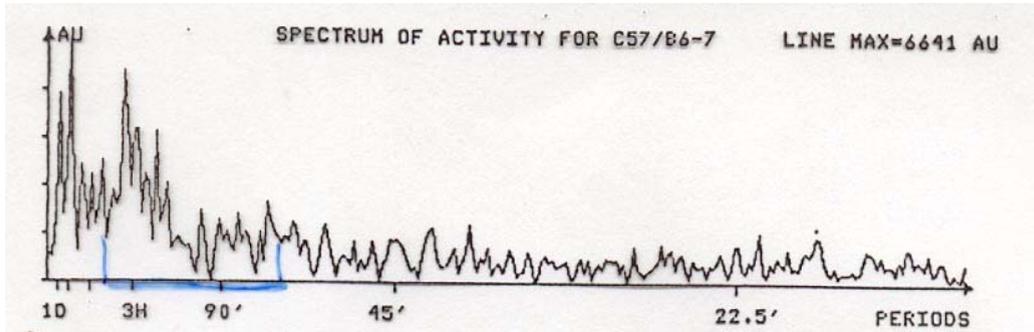
les composantes spectrales

- harmoniques : la forme de la fonction périodique
- composantes non harmoniques : les fréquences des oscillateurs

spectrogramme (plus classiquement spectre)

- il fait apparaître les diverses composantes (sinusoïdes) constituant le signal
- il donne la répartition énergétique de ces composantes en fonction de la fréquence ; *attention* $f = 1 / T$
 - composantes ultradiennes
 - composantes circadiennes
 - composantes infradiennes

analyse spectrale (exemple)



$$W = 72 \text{ heures}$$

$$T_e = 8,44 \text{ min}$$

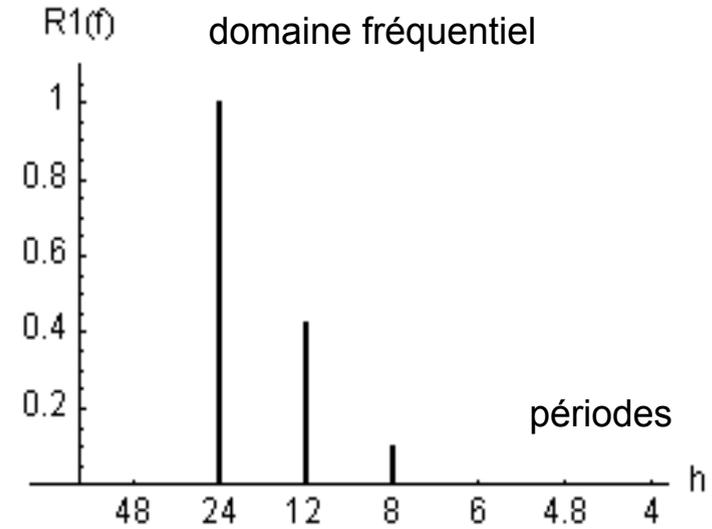
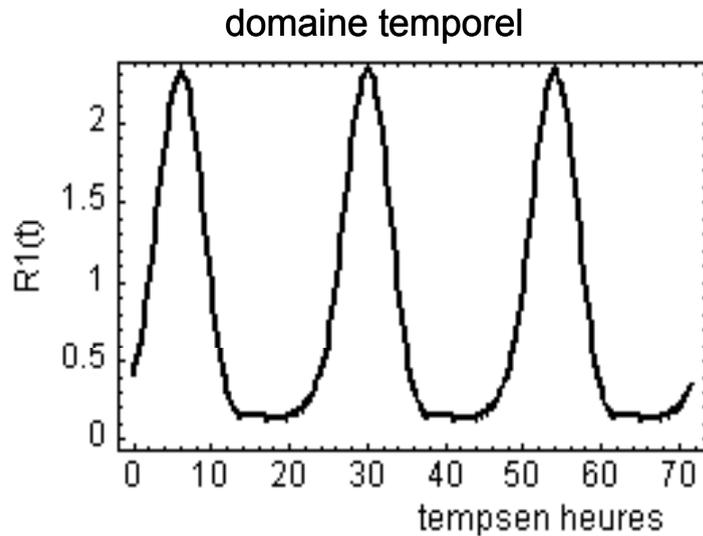
$$\rightarrow F_{\max} = 1/16.9$$

$$N = 512$$

zones spectrales utiles

relation entre domaines fréquentiel et temporel - 1

exemple d'analyse spectrale d'un rythme circadien ; le problème de la raie à 12 h



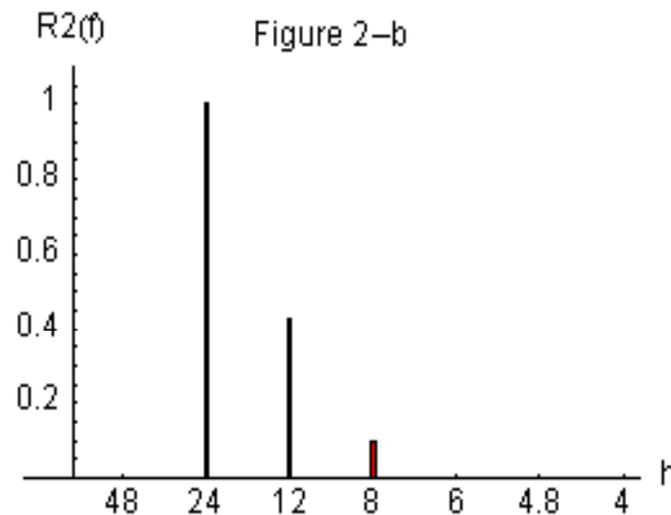
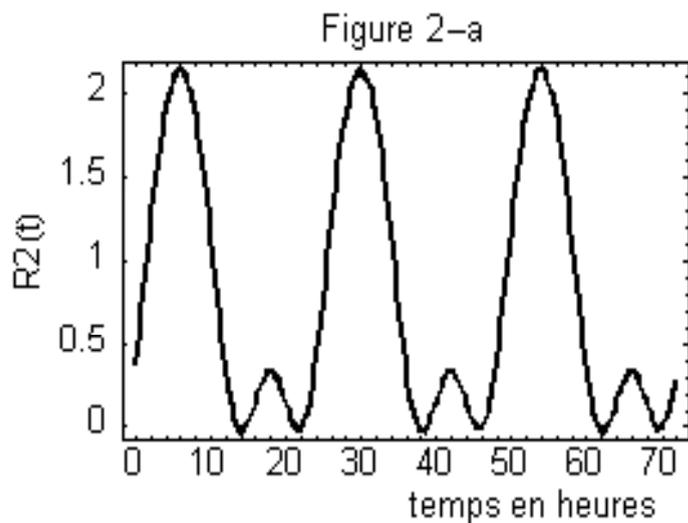
représentation par une série de Fourier $R(t) = \sum_{i=0}^{\infty} a_i \sin(i\omega t + \varphi_i)$, avec $\omega = 2\pi / T$

$$R1(t) = \sin(\omega t) + 0.43 \sin(2\omega t - 1.65) + 0.1 \sin(3\omega t + 2.8)$$

pas de composante bimodale à 12 heures visible sur la structure temporelle bien que le spectre montre une raie harmonique à 12 h de 43 %

relation entre domaines fréquentiel et temporel - 2

un deuxième exemple de rythme circadien



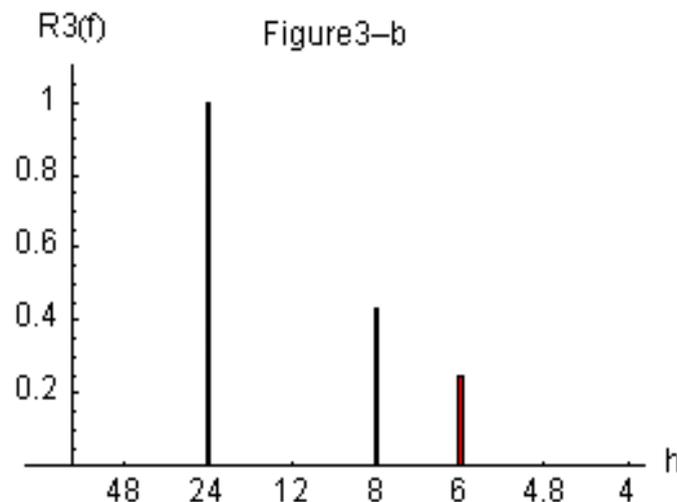
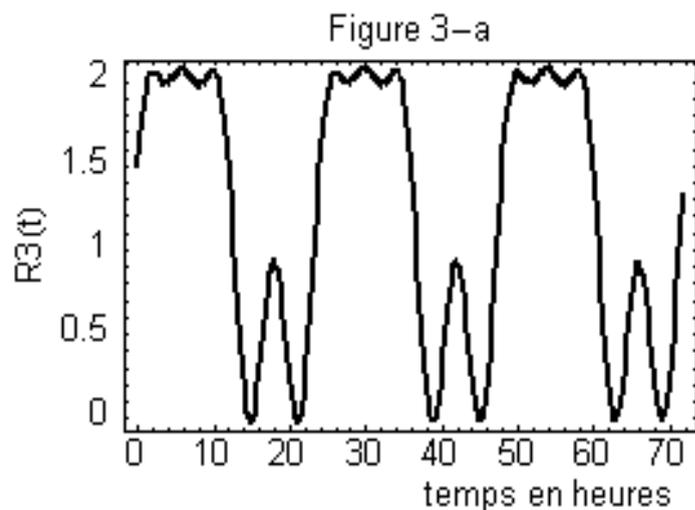
$$R2(t) = \sin(\omega t) + 0.43 \sin(2\omega t - 1.59) + 0.1 \sin(3\omega t)$$

apparition d'une composante bimodale à 12 heures 12.5% visible sur le diagramme temporel avec une raie harmonique à 12 h de 43 %

la différence avec le cas précédent est seulement due aux relations de phases des diverses harmoniques

relation entre domaines fréquentiel et temporel - 3

un troisième exemple de rythme circadien

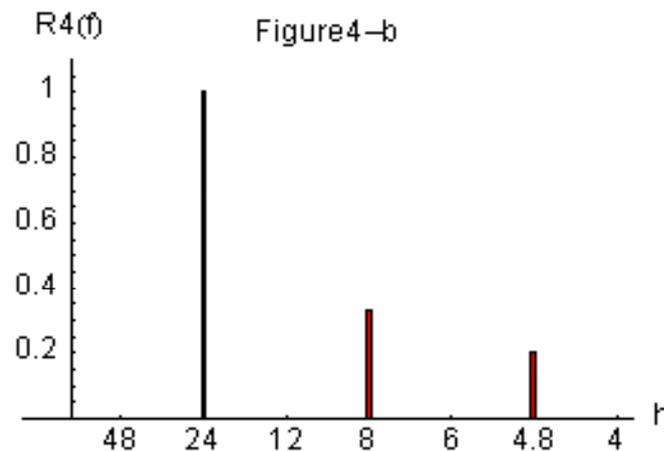
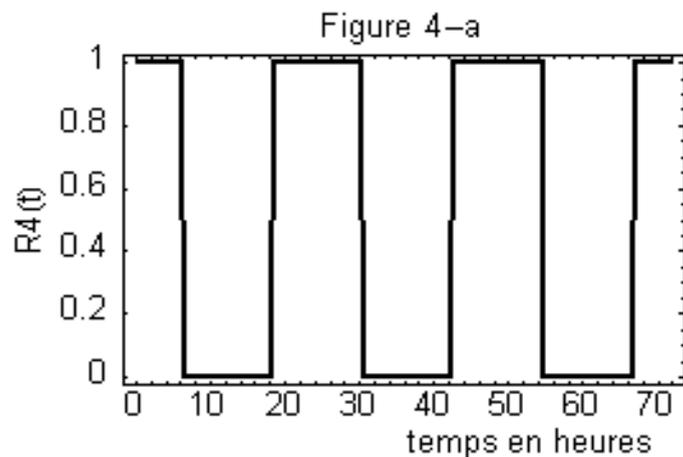


$$R3(t) = \sin(\omega t) + 0.43 \sin(3\omega t) + 0.25 \sin(4\omega t + 1.6)$$

apparition d'une composante bimodale à 12 heures visible sur la structure temporelle bien qu'il n'y ait pas d'harmonique à 12 h !

relation entre domaines fréquentiel et temporel - 4

un dernier exemple de rythme circadien



$$R4(t) = \sum_{i=0}^{\infty} a_i \cos(i\omega t), \text{ avec } a_i = \frac{2}{\pi i} \text{ pour } i \text{ impair et } 0 \text{ pour } i \text{ pair}$$

forme parfaitement circadienne mais absolument non sinusoidale

analyse spectrale, définitions

La période est le temps séparant la répétition d'un même motif ; elle est reliée à la fréquence par :

$$T = \frac{1}{F}$$

24 h → 11.57 μHz

La notion de fréquence évoque la **répétition** dans le temps d'un même motif (par exemple la sinusoïde). On peut distinguer :

- La fréquence au sens de Fourier :

Permet de représenter un signal d'énergie finie par une somme d'exponentielles complexes, lesquelles représentent l'énergie de chaque composante fréquentielle.

$$x(f) = \int_{\mathcal{D}} x(t) \cdot \exp(-2i\pi ft) dt$$

- La fréquence instantanée :

Déterminée à partir de la phase instantanée.

$$f_i = \frac{1}{2\pi} \frac{d\varphi(t)}{dt}$$

Il est possible de rechercher une répétition avec d'autres formes que des sinusoïdes, ou une invariance de formes à des échelles de temps et de fréquences différentes ; on parle alors « d'échelles », ex : les ondelettes.

conclusion : parler de période c'est impliquer la stationnarité

transformation de Fourier par FFT

Le spectre d'un signal est généralement obtenu par calcul de la transformée de Fourier du signal. La transformée de Fourier est généralement calculée à l'aide d'une FFT (*Fast Fourier Transform*).

Le calcul doit s'opérer sur un nombre $N = 2^p$ échantillons. Dans ces conditions, si T_e est la période d'échantillonnage, les échantillons fréquentiels sont donnés tous les $1/N T_e$. N est naturellement lié à T_w (la largeur de la fenêtre) par $N T_e = T_w$

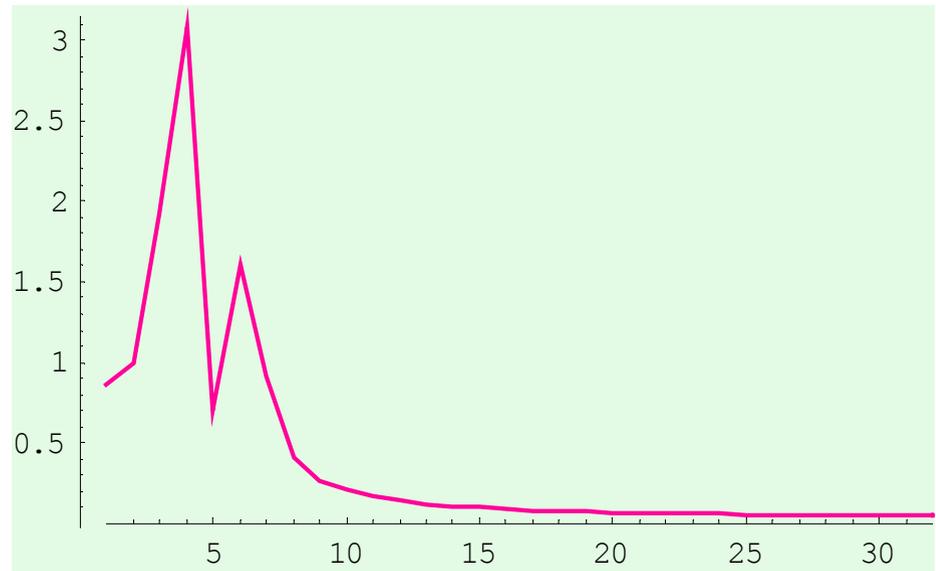
exemple sur un signal simulé :

$$R(t) = 1 * \sin\left(\frac{2\pi}{24}t\right) + 0.5 * \sin\left(\frac{2\pi}{12}t\right)$$

$T_e = 1h$; $N = 64$; $T_w = 2.67j$

seul 32 échantillons fréquentiels sont utiles

la discrétisation fréquentielle donne 64 h, 32 h, 21.33h, 16h, etc.



conclusion : on ne peut pas trouver 24 h !

mais l'ajout de zéros peut pallier les problèmes de faible N.

transformation par FFT : ajout de zéros

Le signal est complété par des zéros de telle sorte que la durée $T_w' = M \cdot T_w$ sur laquelle sera calculé le spectre, soit beaucoup plus grande que T_w .
 M est également de la forme $M = 2^p$. Dans ces conditions la résolution spectrale des calculs va donner comme pas de fréquence, pour définir le spectre, $1/T_w'$; c'est-à-dire M fois plus de points.

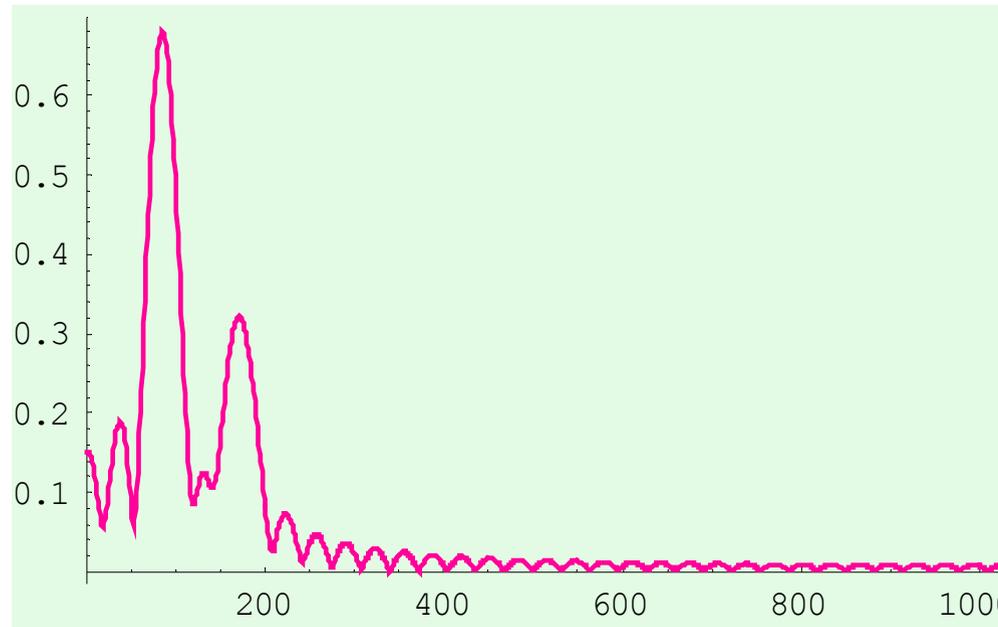
$T_e = 1\text{h}$; $N = 64$; $T_w = 2.67\text{j}$

$M = 32$

la discrétisation fréquentielle donne
autour de 24 h :

24.38 h, 24.09 h, 23.81 h, etc.

**conclusion : on va trouver une
valeur proche de 24 h pour la
raie principale**



***un nouvelle difficulté apparaît : des rebonds parasites semblent être
des raies supplémentaires, en fait elles n'existent pas !
la notion de fenêtre s'impose***

FFT : fenêtres de pondération - 1

En réalité le signal évalué par la FFT est une composition (un produit) du signal étudié et d'une fenêtre d'observation : une fenêtre rectangulaire.

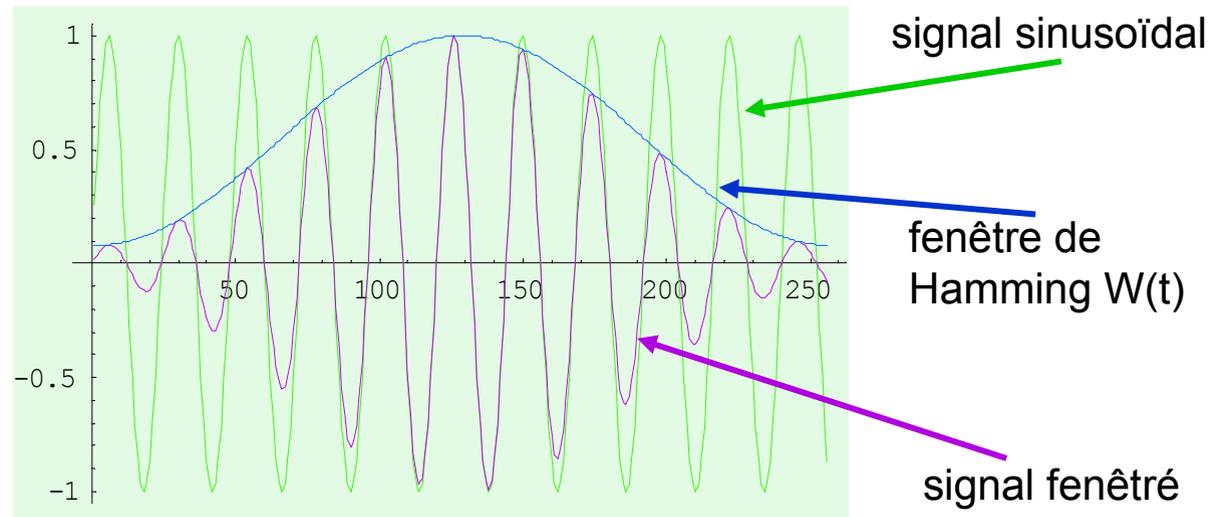
Le spectre calculé est donc une «composition» (une convolution) des spectres du signal (seule partie utile) et de la fenêtre ... dont on n'a que faire.

La fenêtre rectangulaire (naturelle) présente de forts rebonds ; d'autres formes permettent de les atténuer, malheureusement au détriment d'une perte de résolution.

fenêtre de Hamming

$$W(t) = 0.54 - 0.46 * \text{Cos} \left(\frac{1}{T} * \pi * (t-1) \right)$$

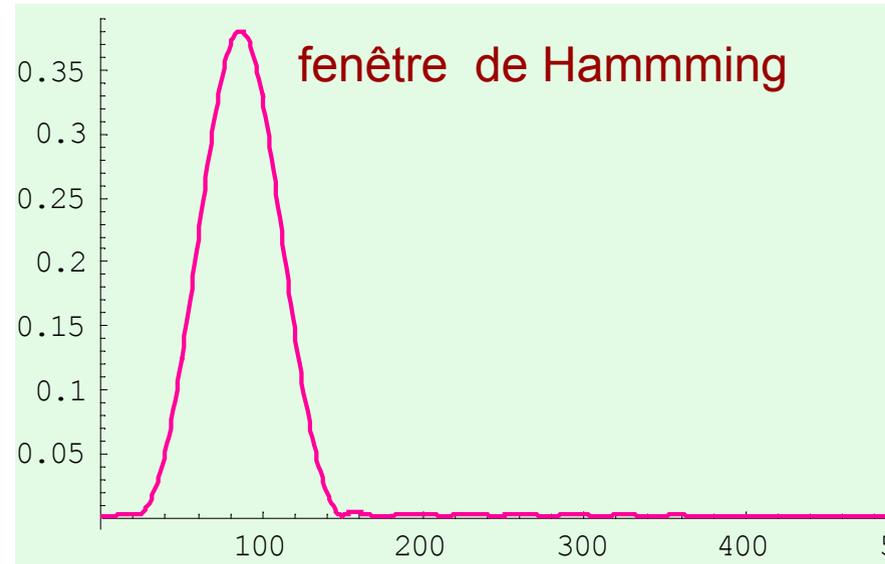
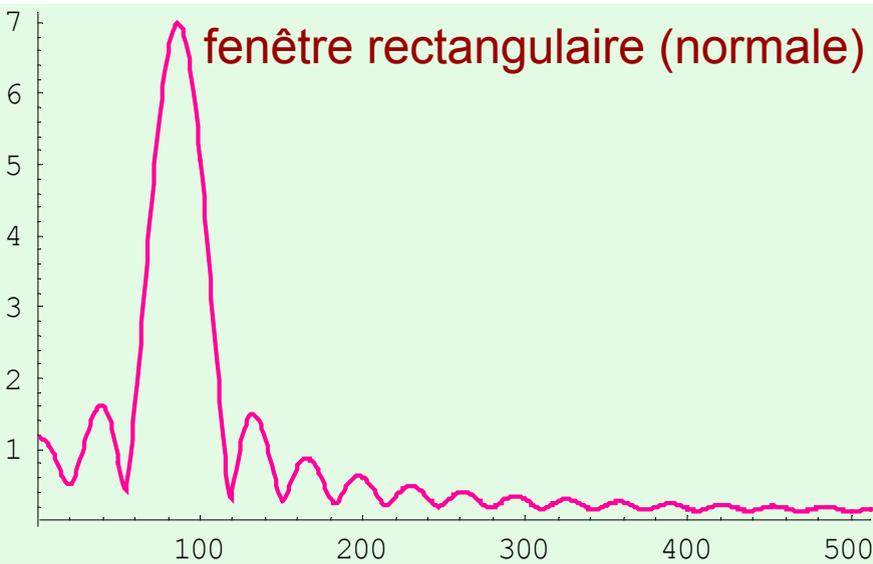
effet temporel d'un fenêtrage sur un signal purement sinusoïdal



FFT : fenêtres de pondération - 2

Cas de l'analyse d'une sinusoïde

$$R(t) = 1 \cdot \sin\left(\frac{2\pi}{24}t\right)$$

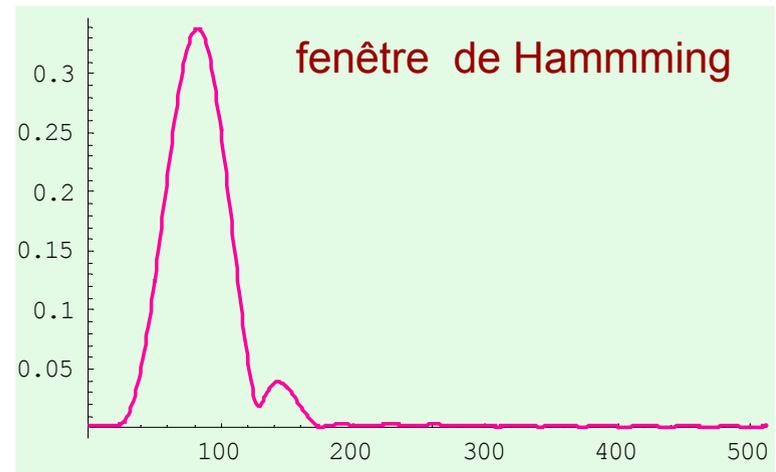
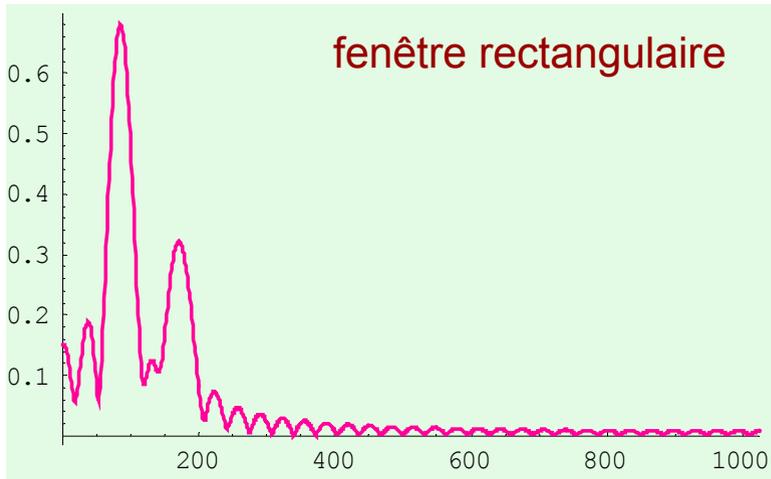


conclusion : le fenêtrage élimine les rebonds, risques d'apparentes raies supplémentaires

une nouvelle difficulté apparaît : la largeur du lobe principal est doublée et les possibilités de discrimination décroissent

FFT : fenêtres de pondération - 3

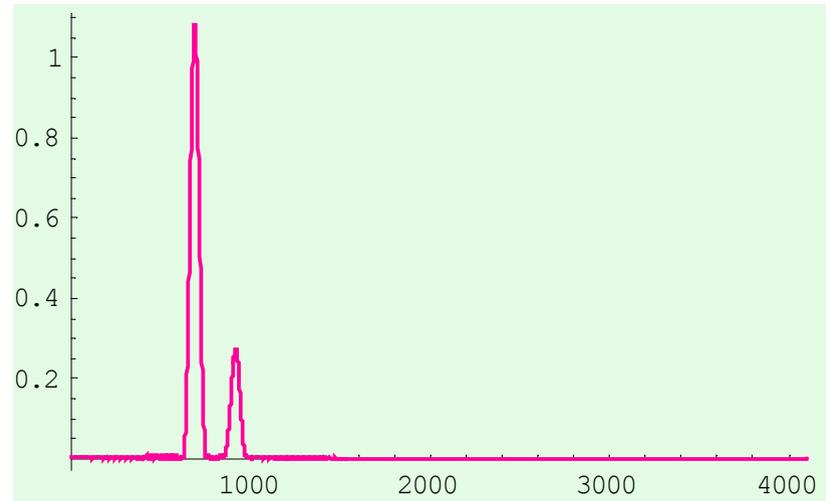
le fenêtrage entraîne une confusion des raies utiles



une seule solution augmenter la largeur de la fenêtre de mesure

utilisation d'une fenêtre de longueur 512 soit 8 fois plus longue (21 jours au lieu de 2.67)

Conclusion : il y a un compromis à faire entre le risque de raies artefactes et celui de la discrimination selon que l'on utilise ou non une fenêtre de pondération ; dans tous les cas plus la longueur de la fenêtre de mesure est grande meilleurs seront les résultats



erreur dans la mesure de la raie circadienne

ous supposons bien sûr être dans un cas synchronisé et donc que la période est 24 h

pourquoi ne trouve-t-on pas 24 h ?

Les divergences potentielles par rapport à 24.000 h :

– **dues au signal lui-même :**

- perturbations du synchroniseur ou compétition de synchroniseurs,
- non stationnarité à court ou moyen terme ;

– **dues au calcul :**

- erreur due à l'échantillonnage du signal et/ou du spectre, ajouts de zéros, fenêtrage,
- erreur due au repliement du spectre,
- erreur due à la présence d'autres composantes du spectre.

une autre question serait aussi : pourquoi on trouve parfois 24 h !

Erreurs liées à la discrétisation du spectre

on suppose que le spectre est déterminé par transformation de Fourier calculée à l'aide d'une FFT.

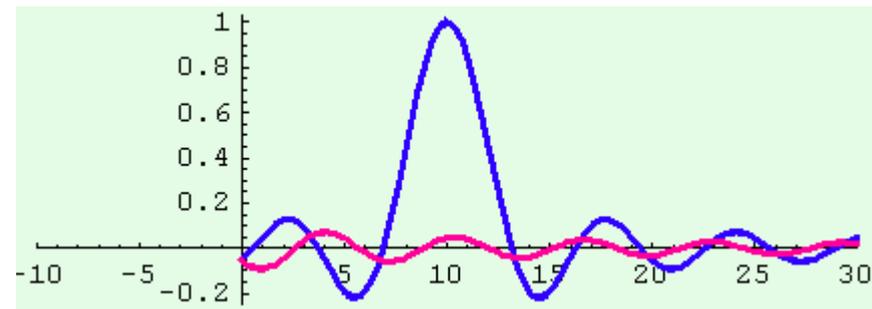
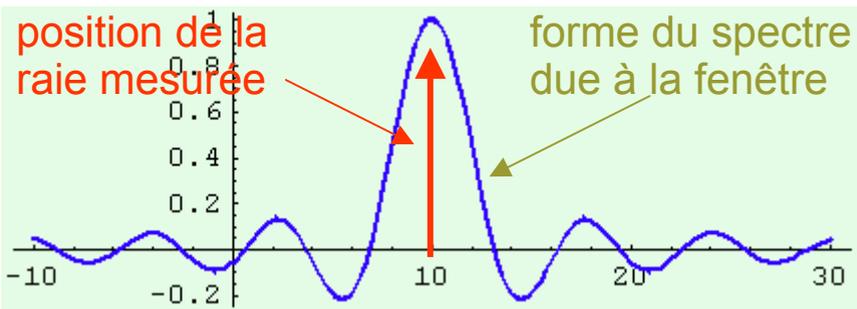
Le calcul doit s'opérer sur un nombre $N = 2^{\text{exp}}$ échantillons. Dans ces conditions, si T_e est la période d'échantillonnage, les échantillons fréquentiels sont donnés tous les $1/N T_e$ ou tous les $1/N * M T_e$ si l'on ajoute des zéros

| exp | $N=2^{\text{exp}}$ | H_0 | H_{sup} | H_{inf} | ΔH_0 | $\Delta H_0/H_0 \%$ | $H_{\text{sup}} - H_{\text{inf}}$ | $H_{\text{sup}} - H_{\text{inf}}/24$ |
|-----------|--------------------|----------------|------------------|------------------|---------------|---------------------|-----------------------------------|--------------------------------------|
| 8 | 256 | 21,3333 | 42,6667 | 14,2222 | -2,6667 | -11,1111 | 28,4444 | 118,5185 |
| 9 | 512 | 21,3333 | 28,4444 | 17,0667 | -2,6667 | -11,1111 | 11,3778 | 47,4074 |
| 10 | 1024 | 24,3810 | 28,4444 | 21,3333 | 0,3810 | 1,5873 | 7,1111 | 29,6296 |
| 11 | 2048 | 24,3810 | 26,2564 | 22,7556 | 0,3810 | 1,5873 | 3,5009 | 14,5863 |
| 12 | 4096 | 24,3810 | 25,2840 | 23,5402 | 0,3810 | 1,5873 | 1,7437 | 7,2655 |
| 13 | 8192 | 23,9532 | 24,3810 | 23,5402 | -0,0468 | -0,1949 | 0,8407 | 3,5030 |
| 14 | 16384 | 23,9532 | 24,1652 | 23,7449 | -0,0468 | -0,1949 | 0,4203 | 1,7511 |
| 15 | 32768 | 23,9532 | 24,0587 | 23,8486 | -0,0468 | -0,1949 | 0,2101 | 0,8755 |
| 16 | 65536 | 24,0059 | 24,0587 | 23,9532 | 0,0059 | 0,0244 | 0,1055 | 0,4391 |
| 17 | 131072 | 24,0059 | 24,0323 | 23,9795 | 0,0059 | 0,0244 | 0,0528 | 0,2196 |
| 18 | 262144 | 24,0059 | 24,0191 | 23,9927 | 0,0059 | 0,0244 | 0,0264 | 0,1098 |
| 19 | 524288 | 23,9993 | 24,0059 | 23,9927 | -0,0007 | -0,0031 | 0,0132 | 0,0549 |
| 20 | 1048576 | 23,9993 | 24,0026 | 23,9960 | -0,0007 | -0,0031 | 0,0066 | 0,0275 |

conclusion : pour obtenir une précision de l'ordre du % 4096 échantillons sont nécessaires, ce que l'on peut obtenir avec un ajout de zéros

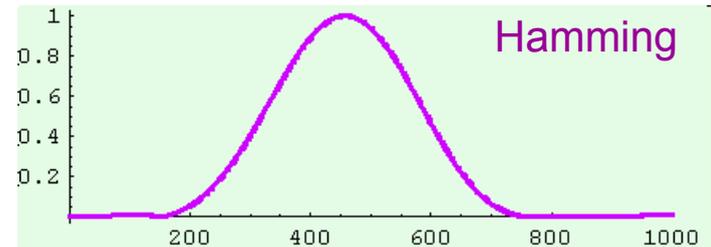
Erreur due au repliement du spectre

repliement du Sinx/x



La somme des raies et du repliement modifie la forme du spectre selon :

- la fréquence de la raie,
- sa phase,
- la largeur et le type de fenêtre.

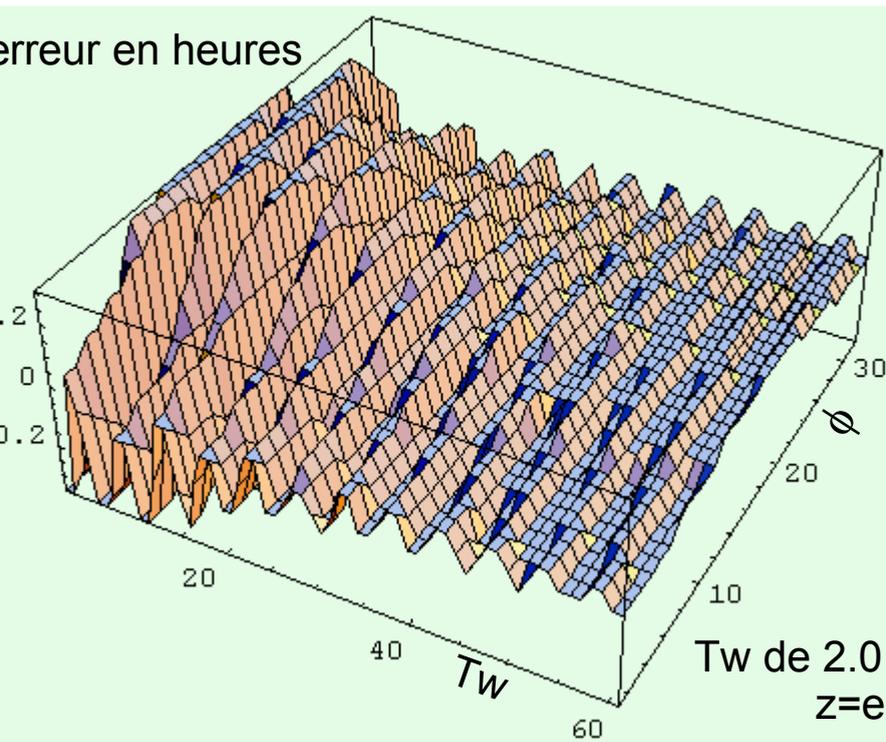


L'étude est effectuée sur une raie circadienne :

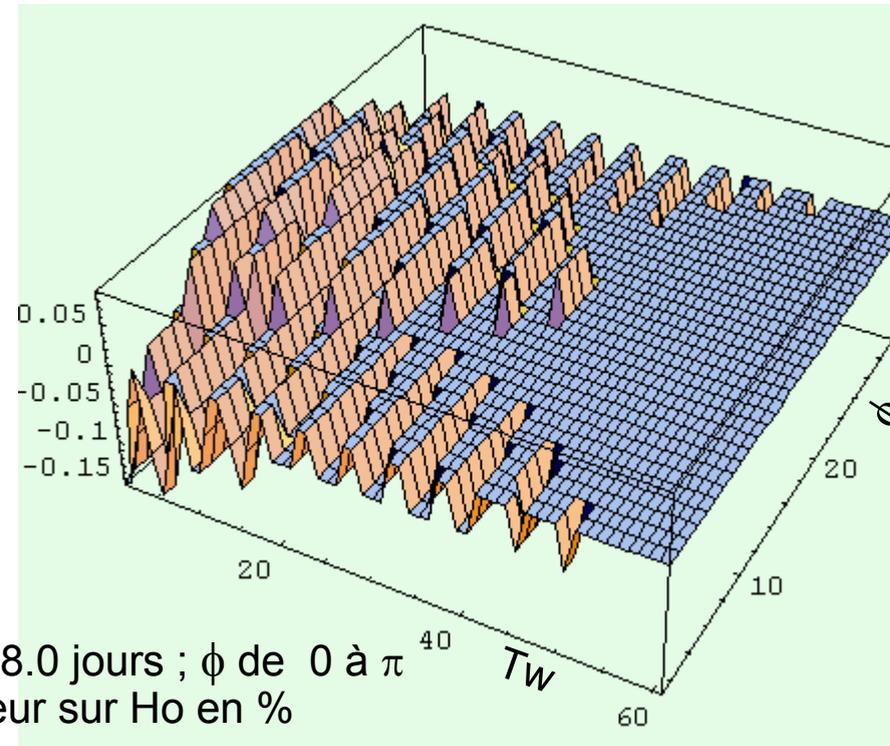
- avec une fenêtre rectangulaire et une fenêtre de Hamming ;
- pour T_w variant de 2.0 à 8.0 jours ;
- et pour ϕ variant de 0 à π ;
- la période et l'amplitude sont étudiées.

évolution de la période

fenêtre rectangulaire

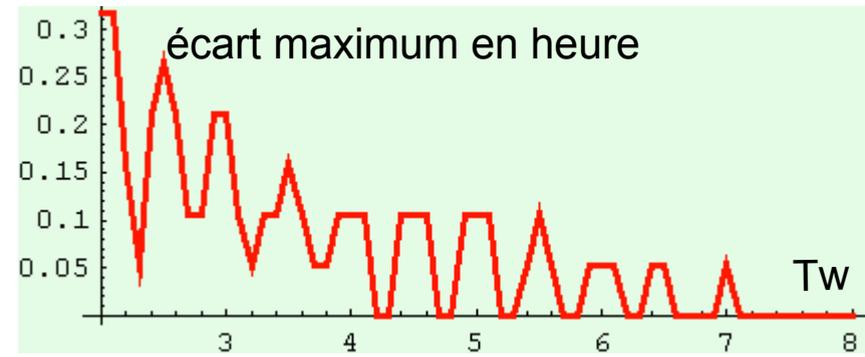
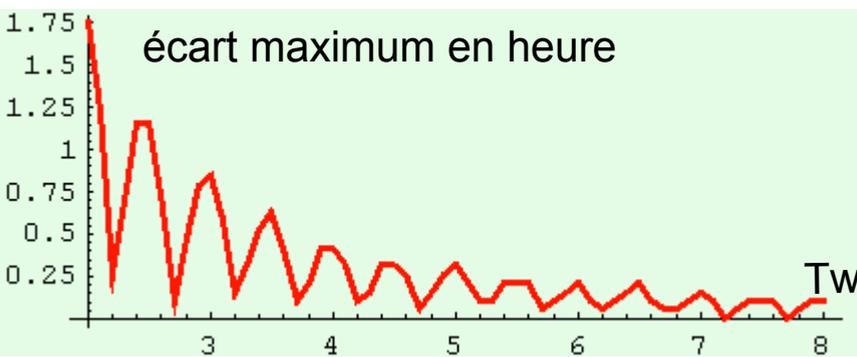


fenêtre de Hamming



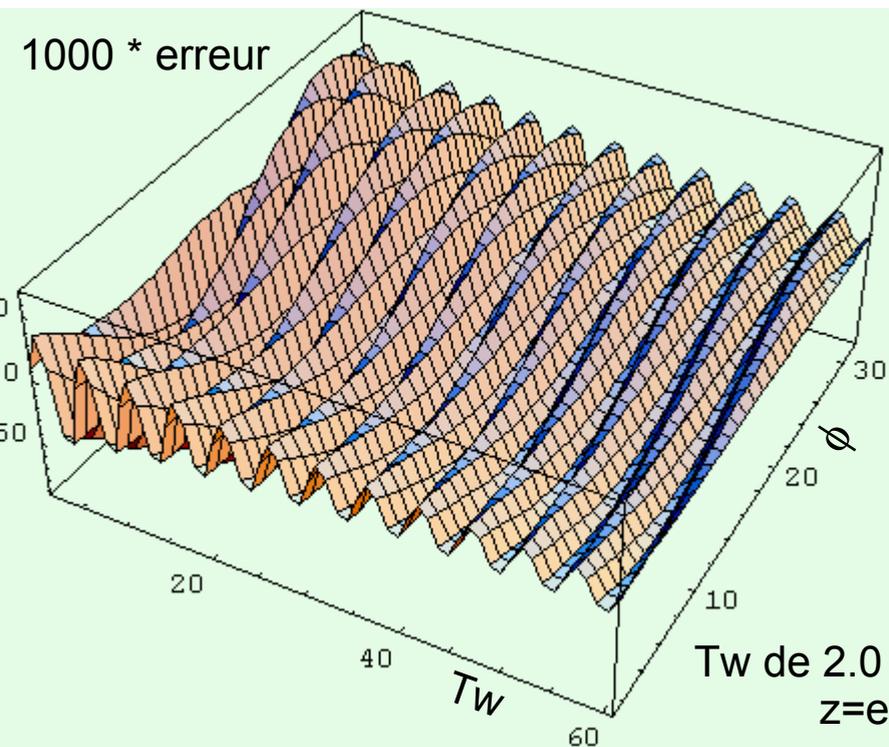
T_w de 2.0 à 8.0 jours ; ϕ de 0 à π

z =erreur sur H_0 en %

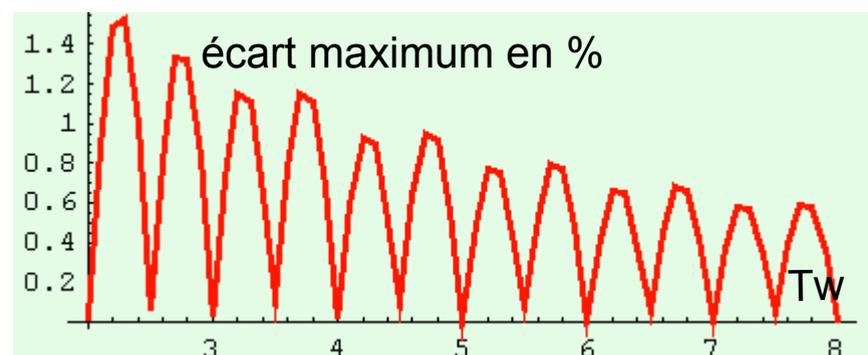
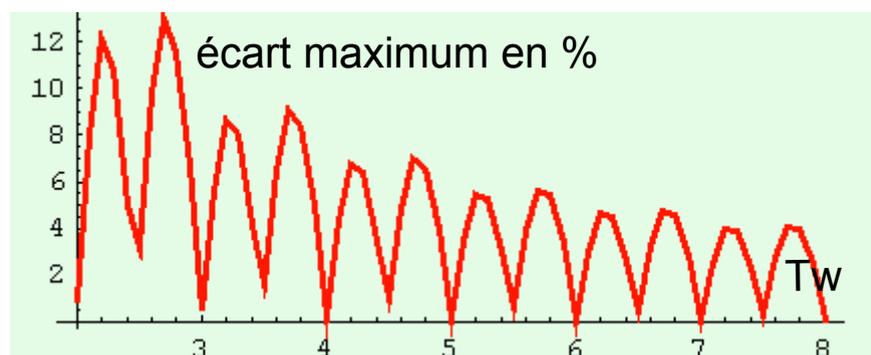
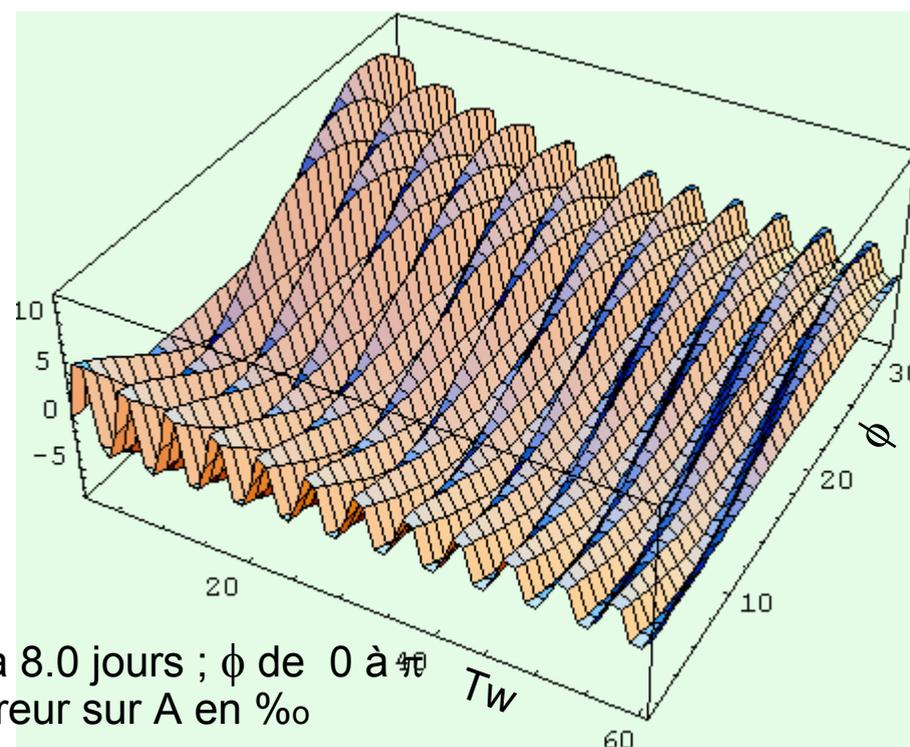


Evolution de l'amplitude

fenêtre rectangulaire

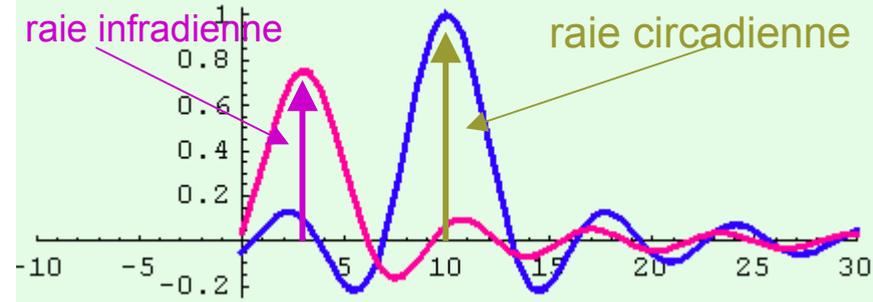


fenêtre de Hamming



erreur due à une composante infradienne

comme pour le repliement du spectre, les composantes infradiennes peuvent perturber la mesure sur les autres raies ... mais beaucoup plus gravement.



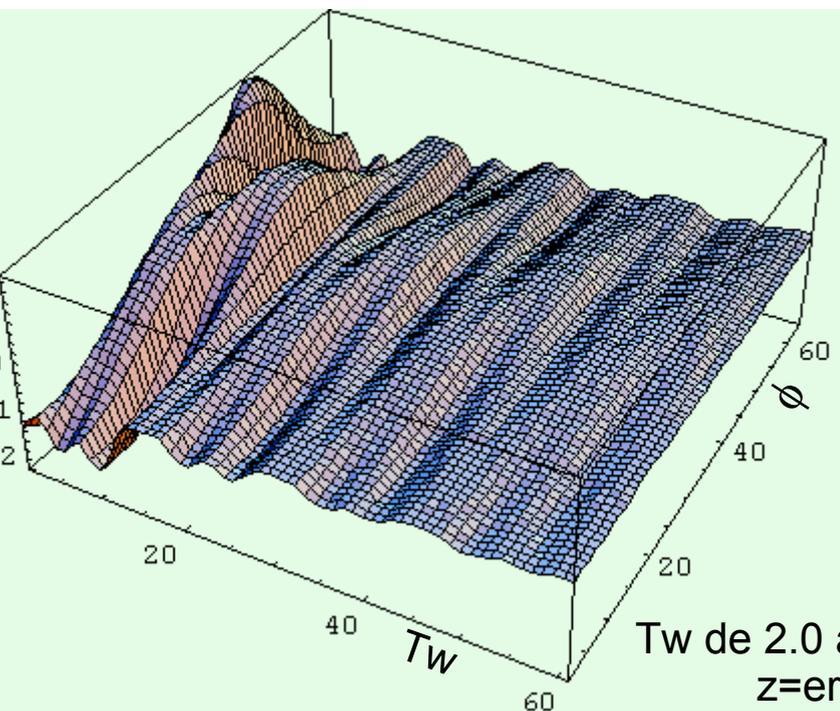
modèle

$$x(t) = \{a \cdot \sin(\omega t) + b \cdot \sin(0.3 \omega t + \phi)\} W(t, T_w) \text{ avec } \omega = \frac{2\pi}{T}$$

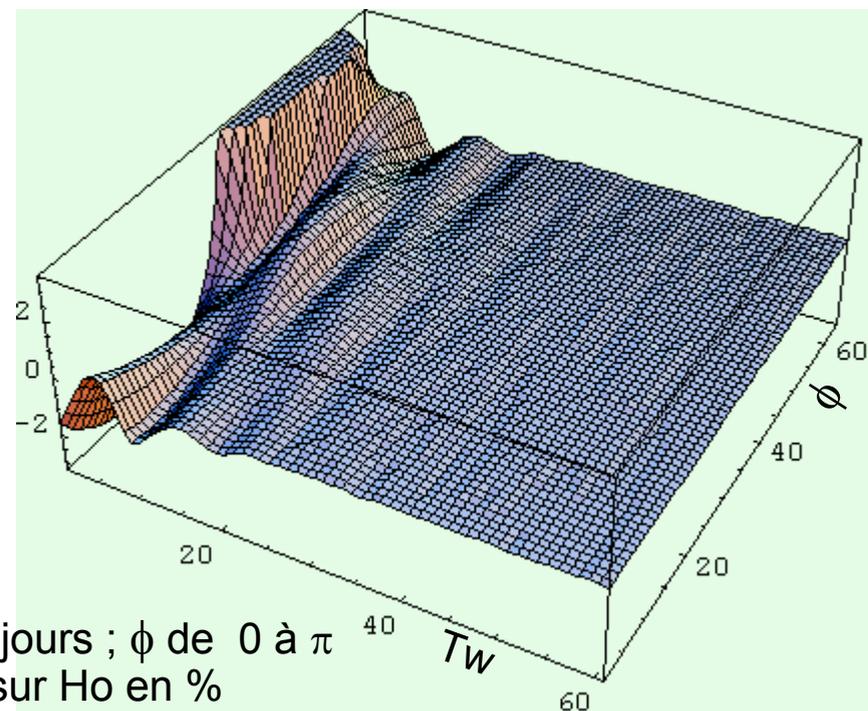
$a = 1.0$ correspond à l'amplitude de la raie circadienne et
 $b = 0.75$ à l'amplitude de la composante infradienne de période 72 h.
 W correspond à la fenêtre de mesure.
 ϕ permet l'étude du phénomène en fonction de la phase relative des deux raies.

évolution de la période

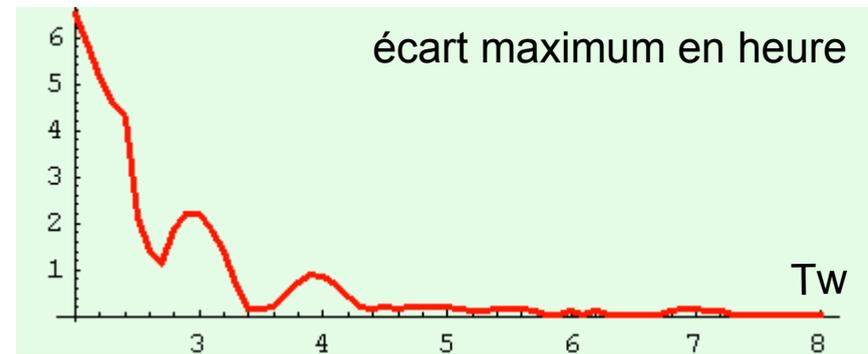
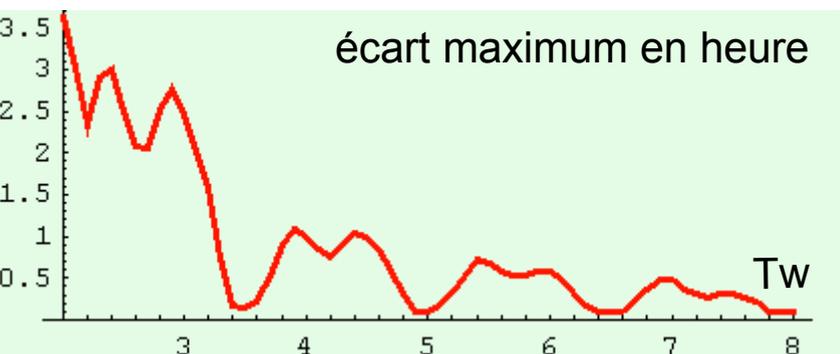
fenêtre rectangulaire



fenêtre de Hamming

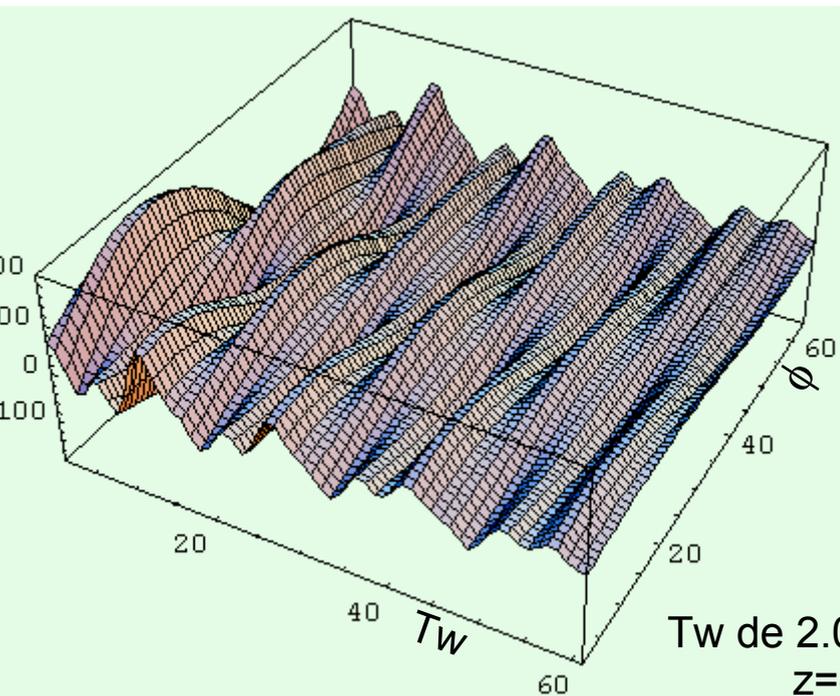


T_w de 2.0 à 8.0 jours ; ϕ de 0 à π
 z =erreur sur H_0 en %

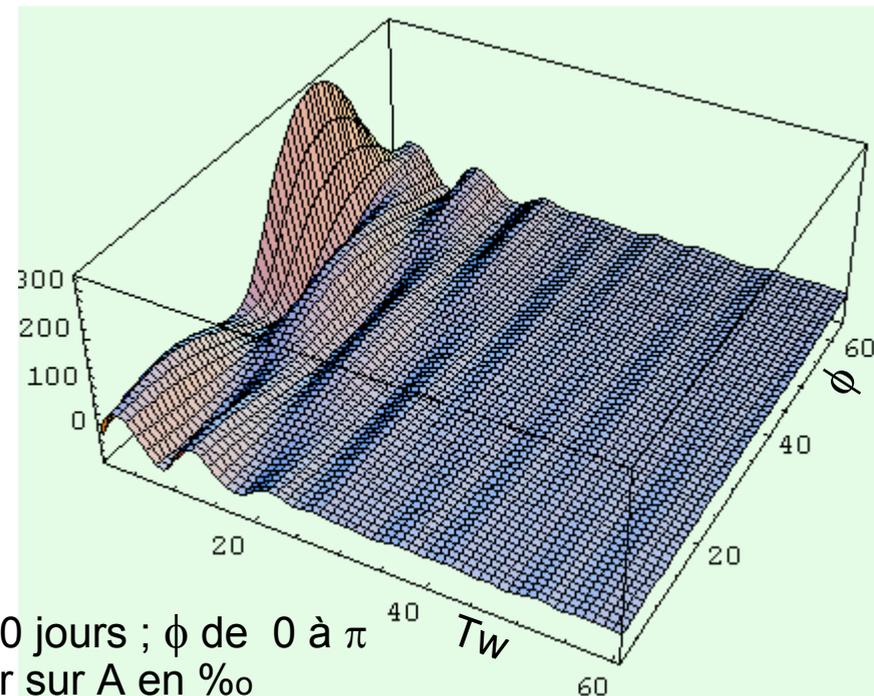


évolution de l'amplitude

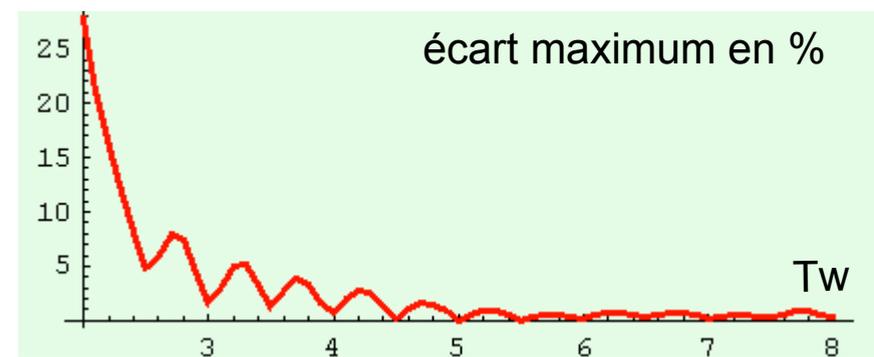
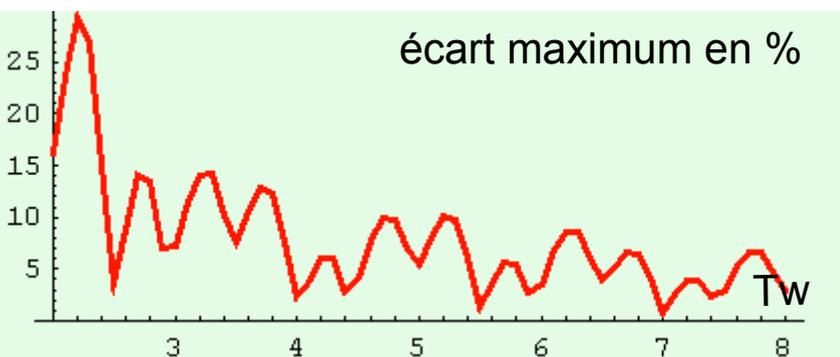
fenêtre rectangulaire



fenêtre de Hamming



T_w de 2.0 à 8.0 jours ; ϕ de 0 à π
 z =erreur sur A en ‰



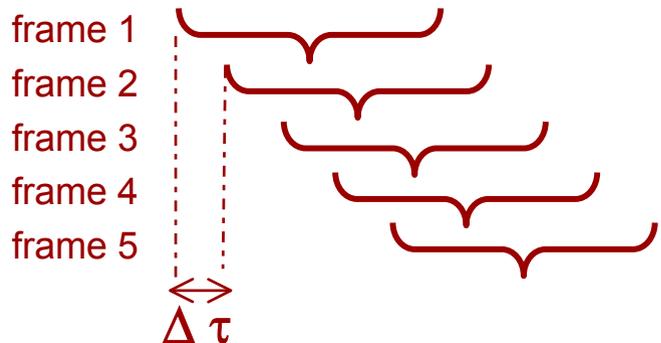
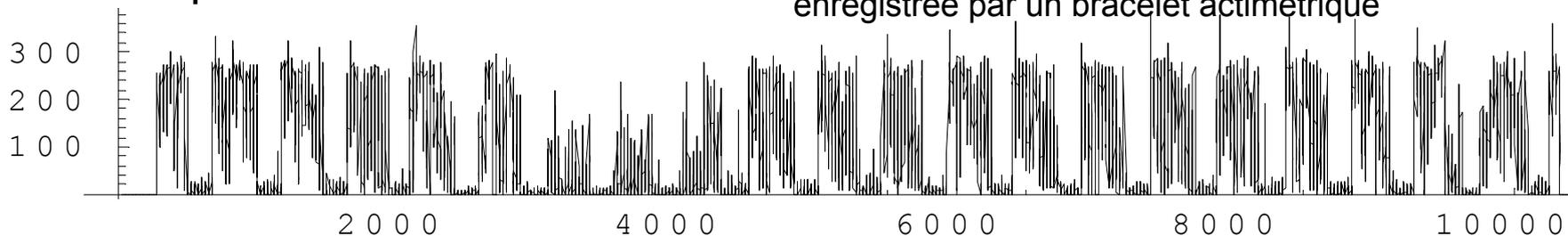
analyse spectrale par fenêtre glissante - 1

l'analyse spectrale globale masque les effets dus à une non stationnarité du signal et ne renseigne donc pas sur son évolution.

l'analyse spectrale glissante est une représentation tridimensionnelle du signal : {amplitude, fréquence, temps}. Elle requiert deux grandeurs T_w , la largeur de la fenêtre, appelée ici frame et $\Delta \tau$ l'écart entre deux frames.

Série temporelle

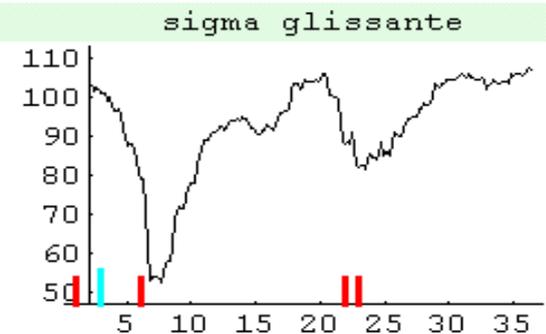
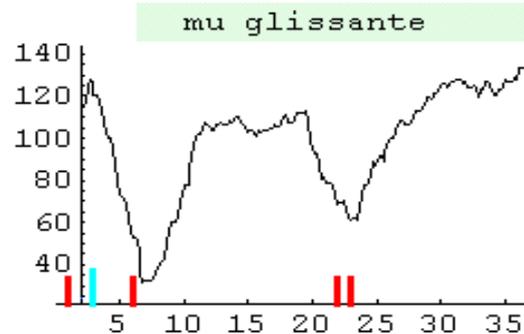
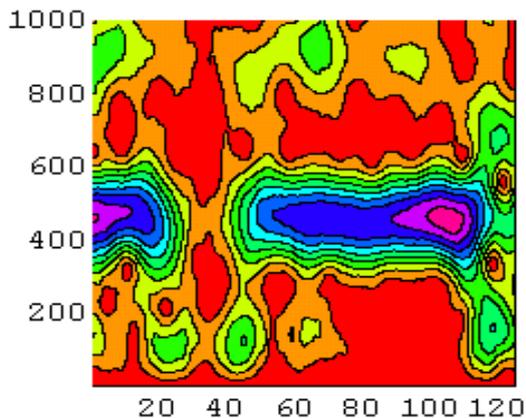
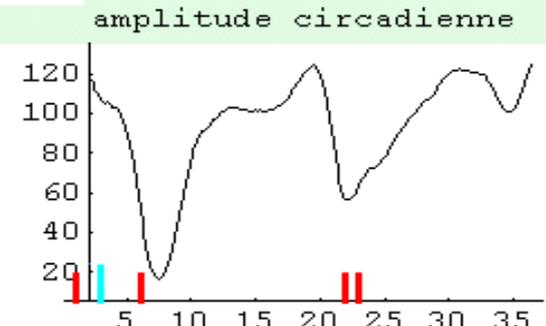
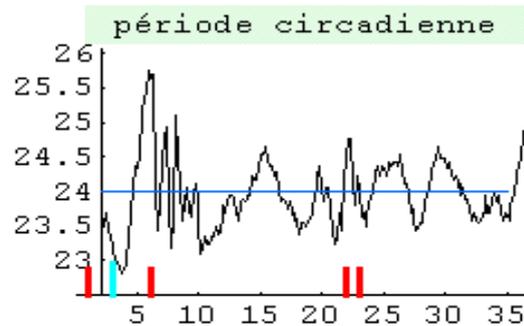
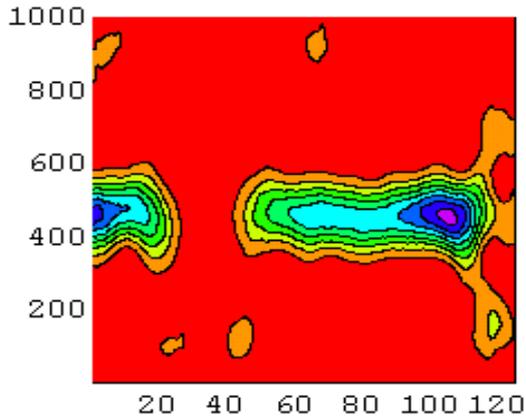
activité d'un patient soumis à une chimiothérapie, enregistrée par un bracelet actimétrique



la représentation tridimensionnelle utilisée est celle des courbes de niveau pour l'amplitude, du n° de frame (le temps) en abscisses et de la période en ordonnées

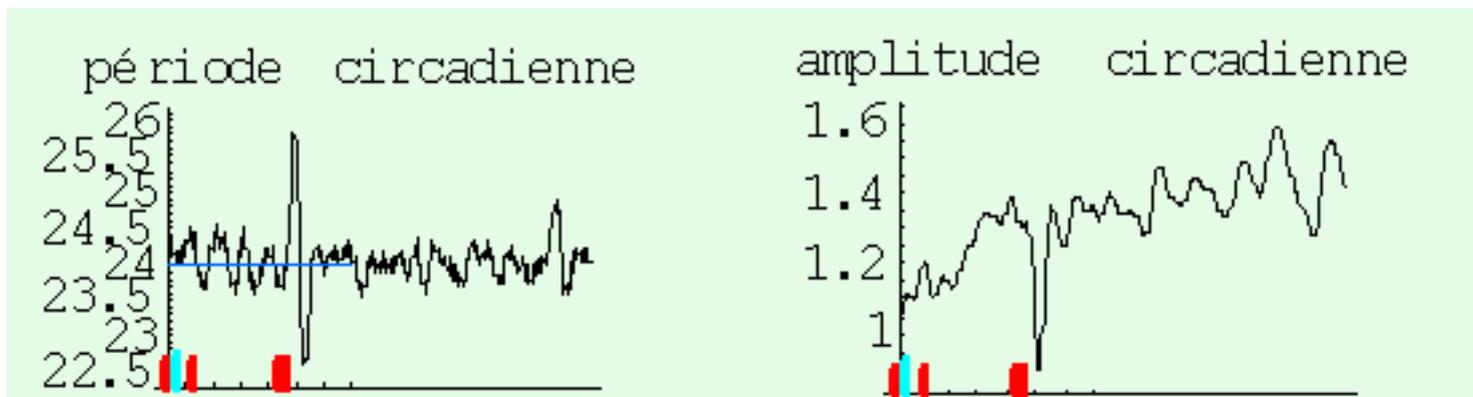
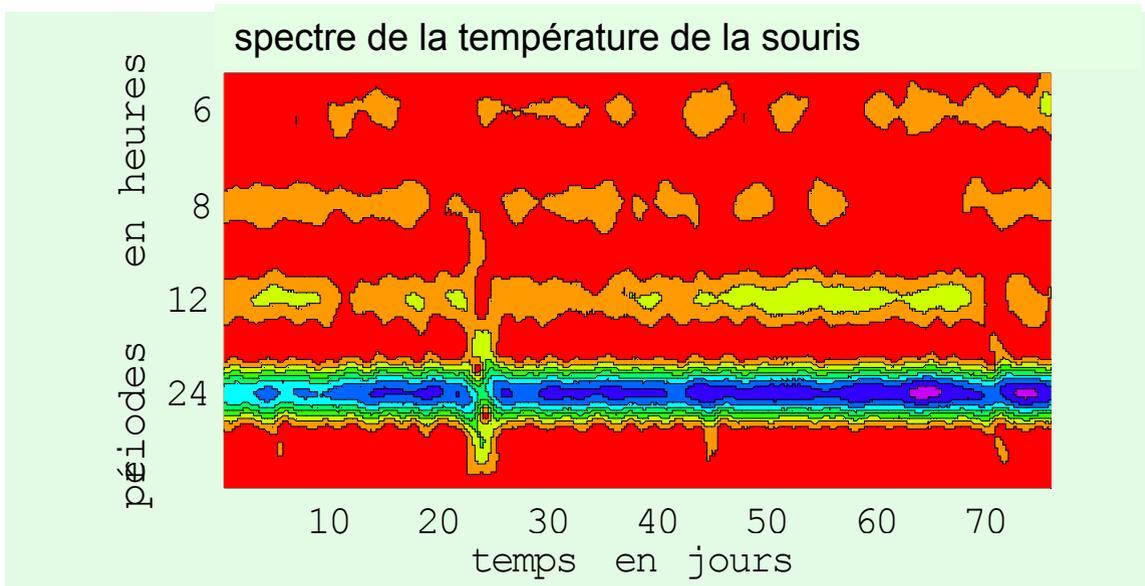
analyse spectrale par fenêtre glissante - 2

- l'analyse spectrale de la série précédente est effectuée avec une frame de 3 jours et un incrément de 4 heures ;
- ces calculs permettent de déterminer : le spectre 3D, les évolutions de T , Amp , μ et σ .



analyse spectrale par fenêtre glissante - 3

autre exemple sur l'évolution de la température enregistrée par télémétrie chez une souris : l'analyse spectrale est effectuée avec une frame de 4 jours et un incrément de 4 heures ;



conclusion

- le mode de saisie des données conditionne les possibilités ultérieures de traitement :
 - la période d'échantillonnage conditionne la période minimale observable,
 - la largeur de la fenêtre d'observation détermine la résolution et le pouvoir séparateur autorisé par l'analyse ;
- l'examen seul du spectre ne peut permettre de conclure sur des éléments relatifs à la morphologie du signal ; il convient toujours de revenir à la structure temporelle ;
- l'outil FFT doit être utilisé avec précaution
 - pour obtenir une bonne précision de calcul il est nécessaire d'utiliser une valeur de $N = 2^p$ suffisante ; $p = 16$ constitue une bonne valeur avec une résolution numérique de 3',
 - si, comme c'est pratiquement toujours le cas, le nombre d'échantillons du signal est trop faible : il suffit d'ajouter des zéros,
 - une fenêtre de pondération peut s'avérer utile mais cela risque de nuire au pouvoir séparateur ;
- afin d'éviter une autoperturbation de la raie circadienne il est souhaitable d'utiliser une fenêtre de largeur T_w de 4 à 5 jours minimum et opérer une pondération de Hamming, la présence d'une raie infradienne (tendance) perturbe considérablement les mesures ;
- l'analyse par fenêtre glissante permet une étude le long du temps pour les phénomènes non stationnaires.