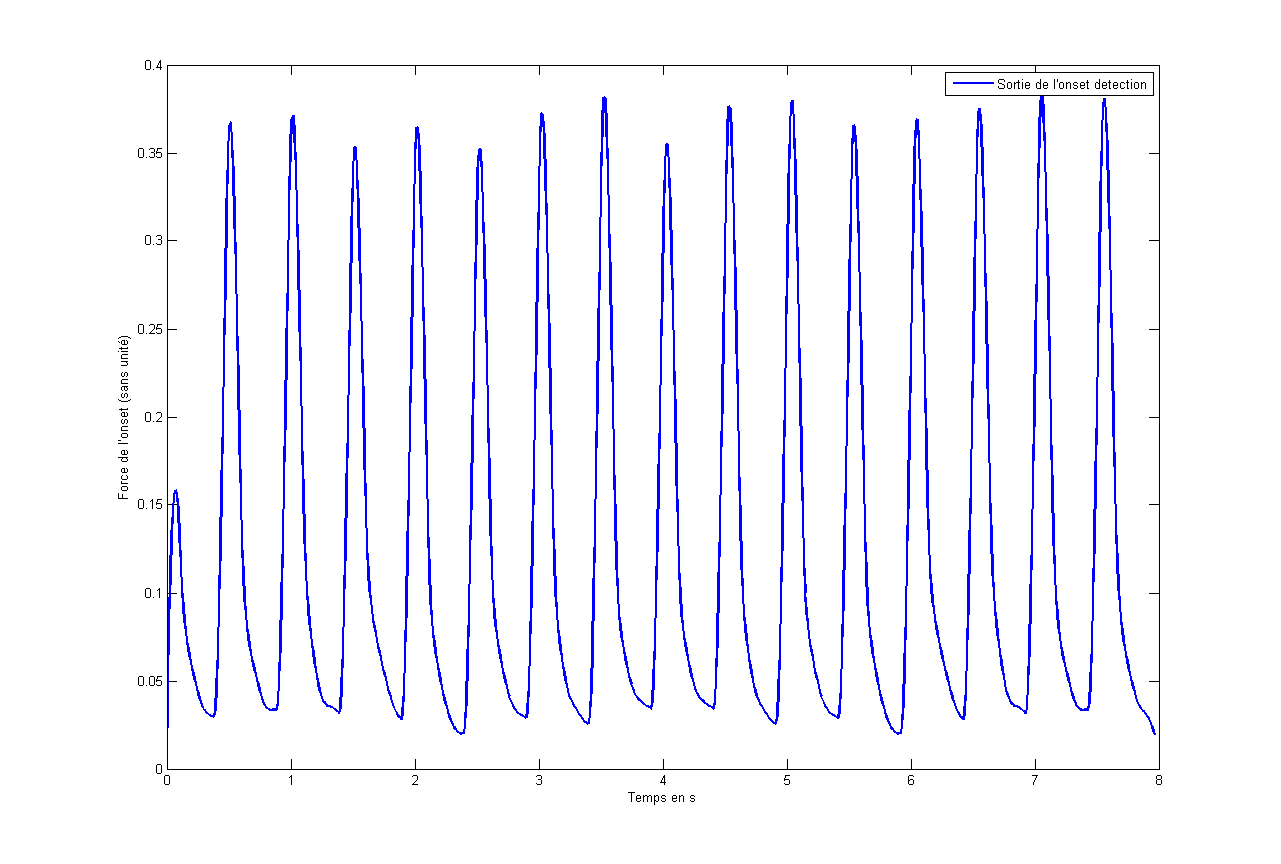
# Détection du tempo

La détection du tempo se fait selon la méthode décrite dans la publication « Beat Tracing by Dynamic Programming » de Ellis. Elle est basée sur l’autocorrélation du signal de sortie d’onset detection.

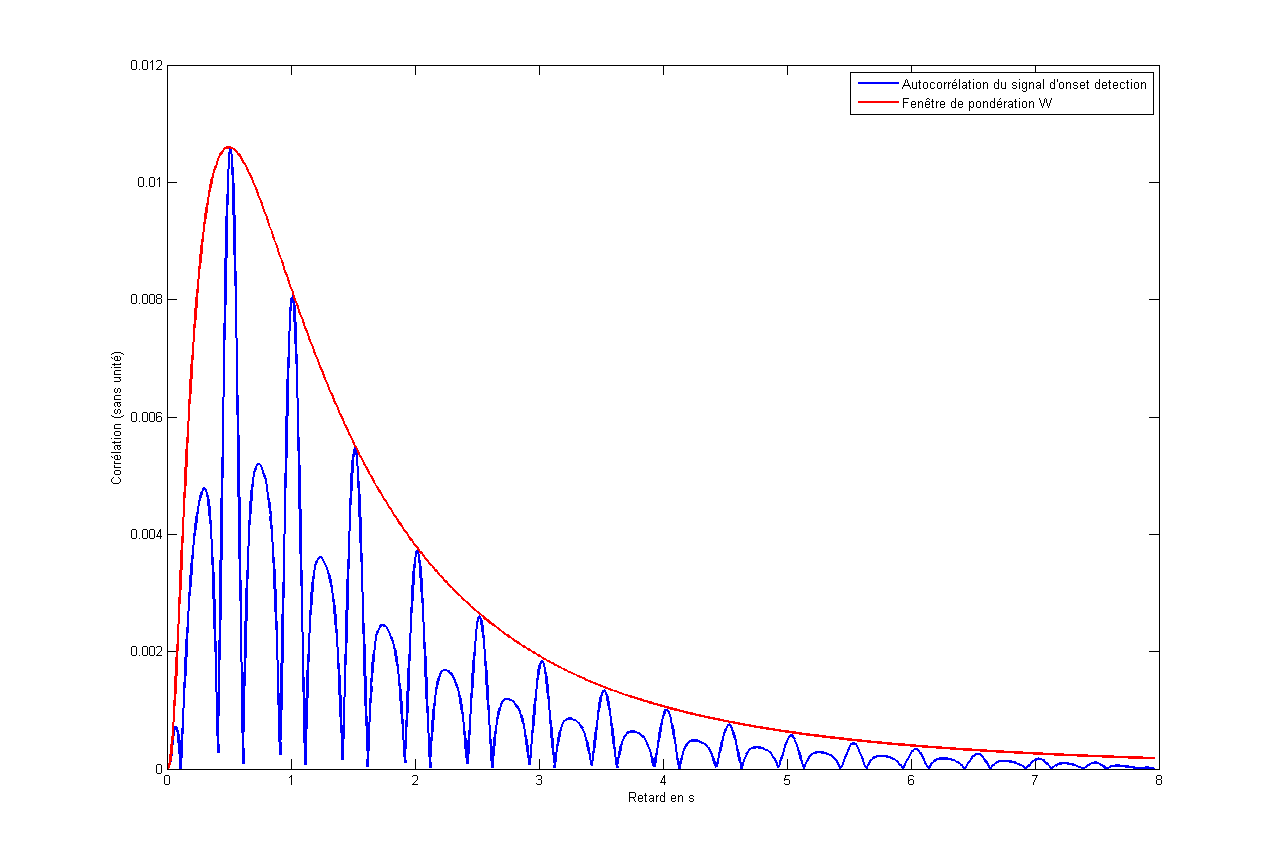
Le principe est que, dans le morceau, le tempo est la conséquence d’une certaine périodicité du signal. Cette périodicité est mise en évidence dans le signal de sortie de l’onset detection :



Dans ce signal, on sent bien qu’il existe une certaine redondance. On met en évidence cette redondance en calculant l’autocorrélation de ce signal. Ellis propose, une fois cette autocorrélation calculée, de l’envelopper par une fenêtre gaussienne décrit ainsi :

Où τ0 est le tempo trouvé communément dans les morceaux musicaux, convertit en seconde. Ellis a effectué un apprentissage de ce paramètre sur la base de données « Beat tracking » du MIREX 2006. La valeur indiquée est de 0.5s soit 120 BPM.

Où στ est la largeur de la fenêtre en octaves. La valeur indiquée est de 1.4 octaves.



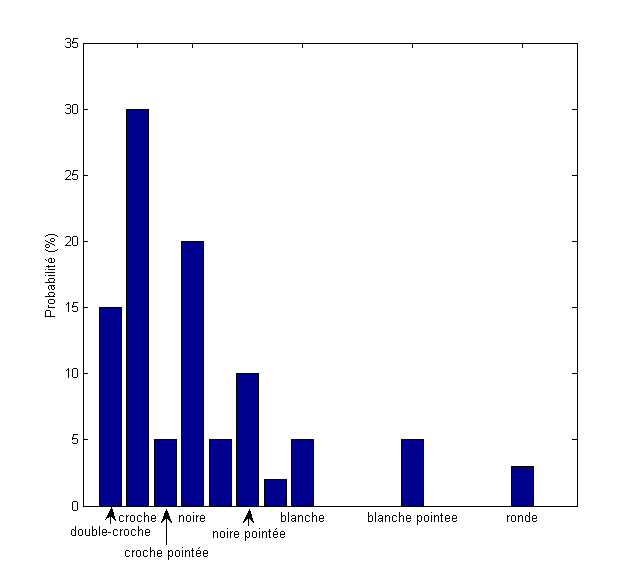
Le tempo du morceau audio est alors celui qui est déduit du retard où la corrélation est maximale. Ici 0.5s donc 120 BPM.

# Détermination de la durée des notes

L’algorithme de détermination de la durée des notes est une classification basée sur l’écart entre deux onsets (deux début de notes), c’est-à-dire la durée en secondes d’une note et sur le tempo déterminé précédemment. Le principe est simple, pour un tempo = 120 BPM, une noire dure environ 0.5s, une croche 0.25s, etc…

Cependant, les durées des notes ne tombent évidemment pas exactement sur des multiples (ou fractions) du tempo. Il faut traiter les durées de notes qui s’en éloigne.

Pour cela, on définit des intervalles pour chaque durée musicale de note (noire, croche, etc…) centrés sur la durée théorique. Cependant, la largeur des intervalles n’est pas linéaire. En effet, il est plus probable de trouver dans un morceau une noire qu’une noire + double-croche par exemple. Si la largeur des intervalles était la même pour ces deux notes, on risque fortement de classer une noire comme une noire + double croche. On choisit donc de rendre la largeur de l’intervalle d’un certain type de note proportionnelle à la probabilité de rencontrer ce type de note.



Les proportions que nous avons fixées sont arbitraires. La « densité de probabilité » se concentre sur les notes courtes (croches et noires) alors que les durée plus longues sont minimisées. Pour certaines durées « étranges », par exemple une blanche + double-croche, nous avons fixée la probabilité à 0 car il est très improbable de rencontrer ce cas de figure.

Nous souhaiterions les vraies probabilités sur plusieurs morceaux pour plus de fiabilité. Nous pensons que ces proportions dépendent fortement du style de musique qui est joué. Un morceau de basse rock contient en général essentiellement des croches et des noires. Un morceau de jazz contiendra en proportion conséquente tous les types de notes.

Il peut encore subsister des erreurs quand une durée de notes est proche de la limite d’un intervalle. Il s’agit d’une situation que l’on peut détecter. Il faudra noter les notes dans ce cas par un marqueur indiquant qu’il est possible que la durée de note soient fausse et qu’ele peut appartenir à la classe directement supérieur (ou inférieure). On utilisera alors c’est marqueur lorsque l’on essaiera de composer des mesures. Si une erreur apparait qui nous empêche de composer une mesure car il y a une incohérence dans la durée des notes, on fera varier les notes qui ont des marqueurs en priorité pour essayer de résoudre le problème.