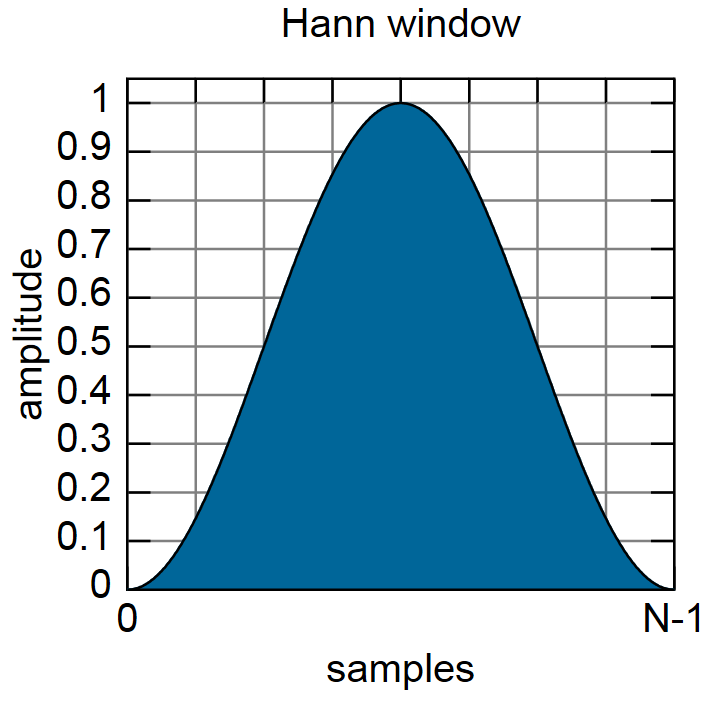
**b) Echantillonnage, fenêtrage**

L’analyse du signal, pour accéder au domaine fréquentiel, s’affranchit de la dépendance temporelle. Le spectre obtenu ne correspond plus à une perception physique, mais à une moyenne temporelle du spectre perçu. Le procédé que nous avons mis en place pour pallier à ce problème est celui le plus couramment utilisé dans ce domaine : l’échantillonnage. Nous avons découpé le signal à traiter en petites séquences, qui, juxtaposées, approximent une échelle temporelle continue.

La taille des échantillons est un paramètre déterminant sur la qualité et la précision de l’analyse combinée finale. Une fois calculé, le spectre ne reflète plus du tout de dépendance temporelle. La durée d’un échantillon correspond ainsi à la durée minimale d’un événement sonore détectable. Il faut donc réduire cette durée autant que possible, pour obtenir une discrétisation temporelle le plus proche possible de la continuité. Il est en revanche nécessaire de conserver un certain nombre de points par échantillons. En effet, le spectre obtenu par l’analyse sera plus précis et proche de la réalité fréquentielle si le nombre de point du signal analysé est important. La meilleure technique pour contourner ce compromis est d’augmenter la fréquence d’échantillonnage. On obtient alors un nombre important de points qui s’étirent peu dans le temps.

Le théorème de Nyquist-Shannon assure qu’un signal reproduit fidèlement toutes les fréquences inférieures à la moitié de sa fréquence d’échantillonnage. Une fréquence d’échantillonnage de 44100Hz (parfois 48000Hz) est donc suffisante pour couvrir la totalité d’une oreille humaine en bonne santé. L’utilisation la plus courante de l’enregistrement audio étant (à notre niveau) la restitution, le matériel et logiciel à notre disposition se cantonnait à ces fréquences d’échantillonnage. Nous avons ainsi dû trouver un compromis entre résolution fréquentielle et précision temporelle. L’hypothèse principale a été que les évènements sonores et variations s’étalant sur une durée inférieure à 20 millisecondes n’étaient pas signifiants pour notre analyse. Le nombre de points a été par cette donnée, couplée à notre fréquence d’échantillonnage lors des enregistrements, à 44100Hz.

 L’échantillonnage introduit par ailleurs des discontinuités aux bornes des morceaux, qui ne sont pas présentes dans le signal original. Le fenêtrage permet de réduire l’effet de ces discontinuités virtuelles. On découpe le signal en plus de morceaux, tout en conservant la même durée pour chaque échantillon. On obtient des "fenêtres", qui se recoupent les unes les autres. Pour que la même partie du signal ne soit pas retraitée à l’identique, on applique une fonction – dite fonction de fenêtrage, ou dans notre cas, fonction de Hann – qui diminue l’importance des valeurs situées aux extrémités de la fenêtre. Ce procédé a le désavantage de démultiplier le temps de calcul des étapes suivantes de l’algorithme (le nombre d’échantillons est bien plus important pour un signal de même longueur). Certaines applications (notamment pour les téléphones portables) devant réduire la complexité au maximum en font donc abstraction. Notre reconnaissance privilégiant plutôt la précision, et disposant d’une puissance de calcul largement suffisante pour conserver un rendu de l’ordre de la seconde, nous avons opté pour un fenêtrage important (recouvrement total d’un échantillon à l’autre), au prix d’une multiplication du temps de calcul par deux.