

LES N-LECTEURS DU GRM PLAYER

Adrien Lefèvre
Ina GRM
alefevre@ina.fr

RÉSUMÉ

Le développement du GRM Player fut, dès le départ, orienté vers l'expérimentation de différents modes de lecture. Outre son architecture client/serveur ou son interface graphique conçue pour des surfaces tactiles multi-point, nous nous intéresserons ici exclusivement à un aspect original du logiciel : sa capacité à confiner plusieurs lectures simultanées dans une même boucle sonore, fixe ou en mouvement.

1. INTRODUCTION

Cette fonctionnalité qui peut sembler surprenante de prime abord, se révèle cependant comme une généralisation de processus électromécaniques ou numériques existants, d'utilisations présentes et passées. Pourtant, le choix d'implémenter des lectures multiples (ou n-lecteurs) dans le Player était au départ purement spéculatif. Par la suite, de nombreux résultats se révélèrent comme des (re)découvertes, souvent inattendues, mais qui s'inscrivent néanmoins dans la longue histoire du traitement et de la perception du signal sonore. Rappelons que le Player fonctionne comme un « magnétophone numérique » [1] et qu'il ne fait que lire le signal sans aucune analyse préalable.



Figure 1 : Phonogène Universel (1962)

Effectuer plusieurs lectures simultanées dans une même boucle, fixe ou en mouvement, permet de nombreux effets sonores. De la démultiplication d'une source lue à différentes vitesses dans une boucle de longue durée (de l'ordre de 10 secondes et plus) créant une rythmique ou quelque ambiance, à la lecture rapide d'une micro-boucle (du millième à quelques centièmes de seconde) générant des timbres lisses ou granuleux, en passant par la mini-boucle (de l'ordre du dixième de seconde) autorisant des transpositions et des étirements-contractions du temps, comme on pouvait le faire il y a cinquante ans avec des têtes de lectures rotatives sur bande, jusqu'à la confection d'un synthétiseur simplifié contrôlé avec Max ; le champ d'application de ces n-lecteurs semble vaste, et réserve encore, j'en suis sûr, de nombreuses surprises.

2. ETUDE DES N-LECTEURS

2.1. Préambule

Avant tout, il semble essentiel de différencier l'unité de mesure du temps réel -la seconde- de l'unité de mesure du temps de défilement d'un lecteur (ou curseur) dans un séquenceur. Cette dernière représente un nombre d'échantillons qui est communément converti et représenté en « secondes », mais il s'agit avant tout d'une *position* ou d'une *distance* exprimée en nombre d'échantillons selon une norme établie (par exemple le standard 44.1 kHz signifie que l'on a 44100 échantillons à lire pour compléter une seconde de temps réel, ou, pour la bande magnétique le 38 cm/s, qui signifie que l'on a 38 centimètres de bande à parcourir pour une seconde de temps réel).

Pour lever toute ambiguïté, nous distinguerons l'unité des secondes réelles, notées *sr*, de l'unité des secondes d'échantillon, notées *se*. Ainsi, une *vitesse de lecture* est mesurée en secondes d'échantillon par secondes réelles (*se/sr*). Pour simplifier, une vitesse peut aussi être considérée comme un coefficient sans unité, qui transforme les secondes réelles en secondes d'échantillon. De même, on emploiera indifféremment les termes « durée » ou « distance », ainsi que « instant » ou « position », pour déterminer l'emplacement d'un lecteur (mesuré en secondes d'échantillon) dans un signal.

2.2. Description d'un n-Lecteur

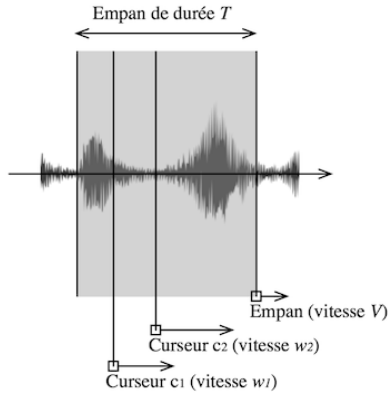


Figure 2 : Représentation d'un 2-Lecteur

Un n-lecteur L est défini par le triplet $L = (T, V, C)$, avec :

- Un *empan de durée T* délimitant la zone temporelle de la boucle [2], en secondes d'échantillon (*se*).
- L'empan pouvant se déplacer à la *vitesse V* dans un signal, en secondes d'échantillon / secondes réelles (*se/sr*).
- Et admettant un ensemble de n curseurs $C = \{c_1, c_2, c_3 \dots c_n\}$. Chaque curseur c_i ($1 \leq i \leq n$) est défini par le triplet $c_i = (p_i, w_i, g_i)$, avec :
 - La *position de départ p_i* dans l'empan à l'initialisation. C'est une valeur normalisée (dans l'intervalle $[0,1]$), et elle n'est pas prise en compte dans la plupart des cas, elle vaut 0 par défaut.
 - Sa *vitesse w_i* exprimée en *se/sr*, indépendante de V .
 - Son gain g_i , avec $g_i \geq 0$.

2.3. Etude des 1-Lecteurs

Que se passe-t-il avec un 1-lecteur quand on fait se mouvoir son empan ? Avec les définitions précédentes, on considère l'empan de durée T , se déplaçant à la vitesse V , et en son sein un curseur partant du début de l'empan se déplace à la vitesse w . Première question : si $w > V$, à quel instant τ le curseur rejoint-il la fin de l'empan (sa partie droite dans la représentation graphique) ?

Soient les trois fonctions :

$$\begin{aligned} C(t) &= w.t && \text{La position du curseur,} \\ E_o(t) &= V.t && \text{La position du début de l'empan,} \\ E_i(t) &= V.t + T && \text{La position de la fin de l'empan.} \end{aligned}$$

$$\text{Pour : } \tau = T / (w - V) \quad \text{On a : } C(\tau) = E_i(\tau) \quad (1)$$

Pour $w < V$, on considère $C'(t) = w.t + T$ (car dans ce cas le curseur part de la fin de l'empan et en rejoint le début) :

$$\text{Pour : } \tau = T / (V - w) \quad \text{On a : } C'(\tau) = E_o(\tau) \quad (2)$$

Dans les deux cas τ est positif. Finalement τ est toujours positif et devient :

$$\tau = T / |w - V| \quad \text{pour tout } w \neq V \text{ et } V > 0. \quad (3)$$

τ est la période du curseur dans le 1-lecteur considéré.

En observant que $w-V$ est la *vitesse relative* du curseur par rapport à l'empan, et qu'il lui faut parcourir la durée T à cette vitesse, on trouve que la position du curseur en fonction du temps réel peut s'exprimer par la fonction affine par morceaux R (que l'on pourrait appeler la *fonction de redistribution du temps*, voir les figures 4 et 5 en fin de document) :

$$R(t) = ([(w - V) \cdot t] \bmod T) + V \cdot t \quad (4)$$

- Pour $w \neq V$ et tout entier $k > 0$: il vient immédiatement :

$$R(t + k \cdot \tau) = R(t) + k \cdot V \cdot \tau \quad (5)$$

- La fonction n'est pas périodique à proprement parler, mais elle s'y apparente et sa période est τ en abscisse et $V \cdot \tau$ en ordonnée.

- Toutes les τ secondes réelles, le curseur se sera déplacé de $V \cdot \tau$ secondes d'échantillon (c'est le *pas d'avancement sur la source*) et aura parcouru $w \cdot \tau$ secondes d'échantillon (c'est la *taille de la fenêtre sur la source*).

- La transposition se fait dans le rapport $w \cdot \tau / \tau = w$ (taille de la fenêtre sur la source / taille de la fenêtre en sortie), elle dépend donc uniquement de la vitesse w , vers le grave ($w < 1$) ou l'aiguë ($w > 1$). La modification de la fréquence perçue dépend de w , mais seulement pour des empan de durée supérieure au dixième de seconde, pour des durées de l'ordre du centième ou moins, la fréquence perçue dépend également de la durée T de l'empan (voir la fin chapitre 3 sur la granulation).

- La dilatation-contraction du temps se fait dans le rapport $V \cdot \tau / \tau = V$ (pas d'avancement sur la source / taille de la fenêtre en sortie), elle dépend donc uniquement de la vitesse V , en dilatation ($V < 1$, la source est répétée par morceaux, voir la figure 4) ou en contraction ($V > 1$, des parties de la source sont ignorées, voir la figure 5).

- On peut caractériser complètement ce système par les trois grandeurs (T, V, w) . Mais tout aussi bien avec $(\tau, V \cdot \tau, w \cdot \tau)$, desquelles on peut déduire les premières. On peut noter que : $|w \cdot \tau - V \cdot \tau| = |w - V| \cdot \tau = T$.

Pour $w = 1$ et $V < 1$, on retrouve le principe de la méthode HOLA (*Homogeneous OverLap-Add*) [3] qui effectue un recouvrement par morceaux répétés de la source vers la sortie. De même, on peut faire le parallèle avec les têtes de lectures rotatives pour bande magnétique fabriquées à la fin des années 1950 par la société allemande Springer [4], il s'agit du même mécanisme : avec V la vitesse de défilement de la bande magnétique, et w la vitesse relative de déplacement de la tête par rapport à la bande (la *vitesse d'exploration* dans la citation ci-dessous).

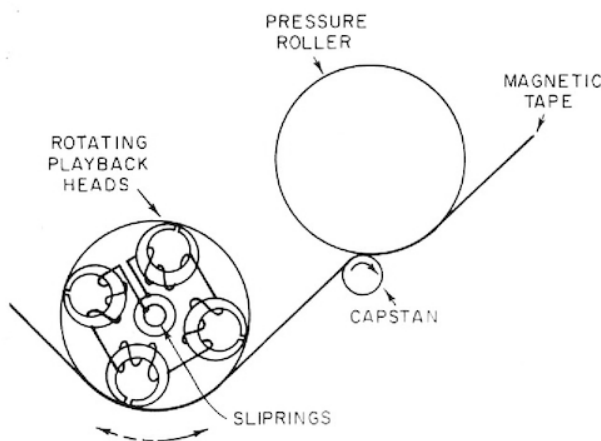


Figure 3. Schematic view of the rotary head assembly (after Springer).

Figure 3 : Tête Rotative Springer (1958)

Voici ce que l'on peut lire dès 1960 dans « La Situation de la Recherche » édité en interne par le Service de la Recherche de la Radio-Télévision Française [5] : *Un appareil de fabrication allemande (Régulateur Temporel) permet de réduire ou d'allonger, dans de larges limites, la durée d'un phénomène sonore sans en altérer la hauteur. Il se compose essentiellement d'un système rotatif de 4 têtes explorant successivement et sans discontinuité une bande enregistrée qui défile devant elles. Pour comprendre le fonctionnement de ce régulateur, il suffit de préciser que la durée de lecture d'un fragment enregistré est fonction de la vitesse de défilement de la bande devant le système de têtes, mais que la hauteur, entendue à la lecture, est fonction de la vitesse d'exploration du signal par le système de tête, c'est-à-dire la vitesse relative de déplacement de la tête par rapport à la bande.*

Puis, deux paragraphes plus loin : *On conçoit qu'une action arbitraire sur les deux vitesses puisse engendrer des transpositions intermédiaires, agissant simultanément sur la hauteur et la durée, dans des rapports déterminés par l'utilisateur. Un appareil de ce type (Phonogène Universel) est actuellement en construction, pour le Groupe de Recherches Musicales et il n'est pas douteux que sa mise en service apporte aux expérimentateurs une nouvelle maîtrise dans la manipulation des objets sonores.*

2.4. Périodicité d'un 2-Lecteur

Comme nous l'avons vu précédemment, la période d'un 1-lecteur est $\tau = T/|w - V|$, et si ce 1-lecteur est fixe ($V = 0$), et son unique curseur est doté d'une vitesse positive ($w > 0$), on a $\tau = T/w$.

Imaginons maintenant un 2-lecteur fixe L , comportant deux curseurs c_1 et c_2 animés de vitesses w_1 et w_2 différentes dans un empan de durée T , et partant tous deux du début de l'empan. Leur période est respectivement $\tau_1 = T/w_1$ et $\tau_2 = T/w_2$. La question est : au bout de combien de temps les deux curseurs vont-ils se retrouver en même temps à leur position de départ, c'est-à-dire tous deux au début de l'empan ?

Par exemple, prenons pour le premier la vitesse $w_1 = 1/2$ et le second la vitesse $w_2 = 1/3$ dans un empan de durée $T = 1$ seconde. Il faudra 2 secondes à c_1 pour parcourir l'empan, et 3 secondes à c_2 pour faire de même. Intuitivement, on comprend qu'il faudra en tout 6 secondes à c_1 et c_2 pour se retrouver en même temps au début de l'empan.

Plus généralement, dans le cas où les périodes τ_1 et τ_2 sont des nombres rationnels (c'est-à-dire pouvant s'exprimer sous forme de fractions d'entiers), il faut considérer le *plus petit commun multiple* [6] des deux périodes τ_1 et τ_2 des deux curseurs de L :

$$Per(L) = ppcm(\tau_1, \tau_2) \quad (6)$$

Où $Per(L)$ désigne la *périodicité* de L . Après quelques équivalences (démonstration 1 en Annexe), on trouve qu'elle est inversement proportionnelle au *plus grand commun diviseur* des vitesses :

$$Per(L) = T / pgcd(w_1, w_2) \quad (7)$$

Par définition, la périodicité d'un n -lecteur mesure la durée au bout de laquelle ce n -lecteur se retrouve dans le même état qu'à son état initial. On remarque que la périodicité est directement proportionnelle à la durée T de l'empan, de plus elle ne change pas si on multiplie par un même facteur la durée de l'empan et les vitesses de curseur. Si les vitesses sont des valeurs entières, le $pgcd$ peut être un entier supérieur à 1, et dans ce cas la périodicité est inférieure à la durée de l'empan.

On observe également que plus il faut de grands nombres pour exprimer précisément les vitesses w_1 et w_2 sous forme de fractions, plus il y a de chances que leur $pgcd$ soit petit (petit numérateur / grand dénominateur), et donc que la périodicité du 2-lecteur soit grande. Lorsque les vitesses ne sont plus dans un rapport rationnel, leur $pgcd$ tend vers zéro, et la périodicité devient infinie.

Toutefois cela n'est jamais vrai dans l'exécution d'un programme d'ordinateur, puisque les nombres qui y sont traités ont une précision limitée et peuvent donc toujours

s'exprimer sous forme de fractions d'entiers. Cependant, l'inverse du *pgcd* pouvant théoriquement atteindre les 18 milliards de milliards (pour un nombre entier standard sur un processeur 64-bits), même avec un empan d'un millième de seconde, on obtient une périodicité de plusieurs centaines de milliards d'années. Un simple 2-lecteur doté de vitesses bien choisies peut donc avoir une périodicité « infinie », en tout cas à notre échelle, et ne jamais se répéter exactement (c'est-à-dire que ses deux curseurs ne se retrouvent jamais simultanément à des positions visitées précédemment).

A l'inverse, un 2-lecteur doté de vitesses s'exprimant sous forme de fractions de petits entiers (2/3, 5/7, 8/19, etc.) aura une périodicité courte. Dans ce cas, les positions de départ des curseurs sont à prendre en compte. Des positions de départ différentes donneront des résultats différents.

2.5. n-Lecteurs à Longue Périodicité

Nous avons vu que la périodicité d'un 2-lecteur était le rapport de la durée de son empan par le *pgcd* des vitesses de ses deux curseurs. Cela est facilement généralisable, et par extension, la périodicité *Per(L)* d'un n-lecteur *L* est donnée par la formule :

$$Per(L) = T / pgcd(w_1, w_2, w_3 \dots w_n) \quad (8)$$

Cette formule reste utilisable pour des vitesses avec un *pgcd* suffisamment grand. Par exemple, si on considère un 3-lecteur *M* avec un empan d'un centième de seconde ($T = 1/100$) doté de trois curseurs avec les vitesses $w_1 = 1.000$, $w_2 = 1.005$ et $w_3 = 1.020$, on obtient une périodicité de 2 secondes :

On a : $w_1=1000/1000$, $w_2=1005/1000$, $w_3=1020/1000$
 Soit : $w_1=200/200$, $w_2=201/200$, $w_3=204/200$
 D'où : $pgcd(w_1, w_2, w_3) = 1/200$
 Donc : $Per(M) = (1/100) / (1/200) = 200/100 = 2 \text{ sec.}$

Mais que se passe-t-il lorsque ce *pgcd* devient trop petit et incidemment la périodicité trop grande, comment malgré tout suivre l'évolution du système ? Lors de l'exécution de micro-boucles avec des empan de durée de l'ordre du centième de seconde ou moins et avec des vitesses appropriées, on constate l'apparition de variations d'amplitude ; une série de « battements » deviennent audibles, entremêlés, lents ou rapides, à peine perceptibles ou manifestes (cela est dû aux déphasages provoqués par la lecture simultanée d'une même portion de signal à des vitesses proches). C'est pour essayer de *visualiser* ce phénomène sonore, même partiellement, que nous allons tenter de *mesurer* l'état d'un n-lecteur par rapport à une référence, au fil du temps.

Pour cela, introduisons la notion de *distance* entre deux instants t_1 et t_2 d'un n-Lecteur *L* : on la note $Dist(L, t_1, t_2)$, et elle correspond à la moyenne quadratique des différences de positions de chaque curseur dans l'empan :

$$Dist(L, t_1, t_2) = \sqrt{\frac{1}{n} \sum_{i=1}^n G_T(w_i t_1, w_i t_2)^2} \quad (9)$$

Où G_T est une fonction de deux variables réelles qui donne la plus courte distance entre deux positions dans l'empan de durée T ; son minimum est 0 lorsque ces deux positions sont confondues et son maximum 1 lorsque ces deux positions diffèrent d'un demi-empan exactement. G_T est une fonction dont les valeurs sont comprises entre 0 et 1, ce qui entraîne que $Dist$ est également une fonction normalisée dont les valeurs sont toujours comprises entre 0 et 1. Après simplification (voir démonstration 2 en Annexe), on trouve :

$$Dist(L, t) = \sqrt{\frac{1}{n} \sum_{i=1}^n H\left(\frac{w_i}{T} t\right)^2} \quad (10)$$

Où : $H(x) = 1 - |1 - 2(x \bmod 1)|$ (11)
 est une fonction périodique « en chapeau ».

Par définition, la fonction $Dist(L)$ est périodique de période $Per(L)$. Lorsque $Per(L)$ n'est pas définie (car infinie), la fonction $Dist(L)$ devient presque-périodique [7].

Sur la figure 6 en fin de document, on peut observer quelques tracés de la fonction $Dist$ avec différents paramètres, l'enveloppe des minimums locaux de la fonction est renforcée en noir, les zéros sont signalés par un petit losange en bas de chaque graphe. Le second zéro correspond à la périodicité, et après chaque zéro la fonction se répète exactement.

3. N-LECTEURS REMARQUABLES

Dans certaines configurations, les n-lecteurs permettent de retrouver ou d'approcher certains effets et filtrages connus. En voici quelques-uns, ainsi qu'une proposition de paramètres pour chacun d'eux :

- Freeze : Un curseur à vitesse normale dans une micro-boucle fixe. Exemple :

$$T = 0.005 \text{ se} / V = 0 / w = 1$$

- Springer (Transposition) : Un curseur à vitesse réduite ou augmentée dans un petit empan à vitesse normale. Exemple :

$$T = 0.05 \text{ se} / V = 1 / w = 0.5$$

- Springer (Time-Stretch) : Un curseur à vitesse normale dans un petit empan à vitesse réduite ou augmentée. Exemple :

$$T = 0.05 \text{ se} / V = 1.2 / w = 1$$

- Multi-Delay : n curseurs avec des positions de départ et des gains décroissants. Exemple :

$$T = 0.5 \text{ se} / V = 1 / n = 4$$

$$C = ((1.0, V, 1.0), (0.9, V, 0.5), (0.8, V, 0.2), (0.7, V, 0.1))$$

- Harmonizer (Multi-Transpositions) : 3 ou 4 curseurs avec des vitesses très différentes dans un empan très court à vitesse normale. Exemple :

$$T = 0.01 \text{ se} / V = 1 / n = 3$$

$$C = ((0, 1, 1), (0, 1.502, 1), (0, 1.997, 1))$$

- Flanger : 2 ou 3 curseurs avec des vitesses perturbées par de l'aléatoire dans un petit empan à vitesse normale. Exemple :

$$T = 0.1 \text{ se} / V = 1 / n = 3$$

$$C = ((0.5, V, 1), (0.5, V+a(t), 0.8), (0.5, V+b(t), 0.7))$$

où a et b sont les fonctions de perturbation aléatoire.

- Multi-Granulation : n curseurs avec des vitesses quelconques (vitesses avec des valeurs proches, vitesses dans des rapports harmoniques, etc.) dans une micro-boucle fixe ou en mouvement (avec un empan à faible vitesse). Le réglage des gains a également un impact direct sur le résultat, les possibilités sont innombrables. Idéalement, mais cela dépend du signal contenu dans la boucle, pour chaque curseur c_i la fréquence fondamentale perçue est w_i / T .

Dans les prochaines versions du Player, il sera possible de choisir le type de fenêtrage appliqué à la micro-boucle (Triangle, Sinus, Hamming, Hann, Gauss).

4. CONCLUSION

Les n-lecteurs ont une structure simple composée d'éléments simples, que l'on peut résumer par : un ou plusieurs curseurs en boucle dans un empan fixe ou en mouvement ; et c'est par la combinaison de ces éléments que l'on obtient des objets complexes. Cette « mécanique » constitue un support pédagogique solide et ludique, tant elle est facile à comprendre. En outre, le moteur audio du Player (NexTape [1]), dont le principal souci de conception fut la qualité du rendu sonore et la précision de la gestion du temps, offre de nouvelles possibilités d'explorer la texture intime du son, et par là même de générer de nouveaux timbres, de nouveaux sons.

Remerciements : à Evelyne Gayou, Daniel Teruggi, Emmanuel Favreau et François Bonnet pour leur relecture et leur contribution à l'élaboration de cet article.

5. ANNEXES

5.1. Démonstration 1

Avec : $\tau_1 = T/w_1$ et $\tau_2 = T/w_2$

On a : $Per(L) = ppcm(\tau_1, \tau_2) = ppcm(T/w_1, T/w_2)$
 $= T \cdot ppcm(1/w_1, 1/w_2)$

Or : $ppcm(1/w_1, 1/w_2) = 1/pgcd(w_1, w_2)$
 Donc : $Per(L) = T/pgcd(w_1, w_2)$

Pour plus d'informations sur le calcul du $pgcd$ et du $ppcm$ de rationnels, voir la page Wikipédia s'y rapportant [6]. On y trouve que :

Avec : $w_1 = p_1/q_1$ et $w_2 = p_2/q_2$ (p_1, q_1, p_2, q_2 entiers)
 On a : $pgcd(w_1, w_2) = pgcd(p_1, p_2) / ppcm(q_1, q_2)$

5.2. Démonstration 2

Soit : $G_T(x_1, x_2) = \frac{2}{T} \min(d, T-d)$

Avec : $d = (x_2 - x_1) \bmod T$

Quels que soient x_1 et x_2 réels et $T > 0$. Où la fonction mod est toujours positive (exemple : $-1.2 \bmod 1 = 0.8$), et où l'on a normalisé $\min(d, T-d)$ par l'inverse de son maximum $T/2$.

Sachant que : $\min(a, b) = \frac{1}{2}(a + b - |a - b|)$

On a : $\min(d, T-d)$

$$\begin{aligned} &= \frac{1}{2}(T - |T - 2d|) = \frac{T}{2} \left(1 - \left| 1 - \frac{2d}{T} \right| \right) \\ &= \frac{T}{2} \left(1 - \left| 1 - \frac{2}{T}((x_2 - x_1) \bmod T) \right| \right) \\ &= \frac{T}{2} \left(1 - \left| 1 - 2 \left(\frac{x_2 - x_1}{T} \bmod 1 \right) \right| \right) \end{aligned}$$

D'où :

$$G_T(x_1, x_2) = \left(1 - \left| 1 - 2 \left(\frac{x_2 - x_1}{T} \bmod 1 \right) \right| \right)$$

Avec : $H(x) = 1 - |1 - 2(x \bmod 1)|$

On a : $G_T(x_1, x_2) = H\left(\frac{x_2 - x_1}{T}\right)$

D'où : $G_T(w_1 t_1, w_1 t_2) = H\left(\frac{w_i}{T}(t_2 - t_1)\right)$

Ou encore, en posant $t = t_2$ et $t_1 = 0$:

$$G_T(w_i t_1, w_i t_2) = H\left(\frac{w_i}{T} t\right)$$

6. REFERENCES

- [1] A. Lefèvre, « Le GRM-Player », *Actes des Journées d'Informatique Musicale 2016*, pp. 170-171, Albi, France, 2016.
- [2] « Empan » : Terme utilisé fort à propos par François Delalande, dans « L'invention musicale », *Actes du 3ème congrès européen d'analyse musicale*, page 4, Montpellier, 2001.
- [3] OverLap-Add Method, https://en.wikipedia.org/wiki/Overlap-add_method (consulté le 7 février 2017).
- [4] Mechanically audio time compression / expansion, <http://www.preservationssound.com/?p=6165> (consulté le 24 janvier 2017).
- [5] Service de la Recherche de la Radio-Télévision Française, dans *Situation de la Recherche 1960*, pp. 237-238.
- [6] PGCD de fractions, https://fr.wikipedia.org/wiki/Plus_grand_commun_diviseur#PGCD_de_fractions (consulté le 10 janvier 2017).
- [7] Fonction presque périodique, https://fr.wikipedia.org/wiki/Fonction_presque_périodique (consulté le 16 janvier 2017).

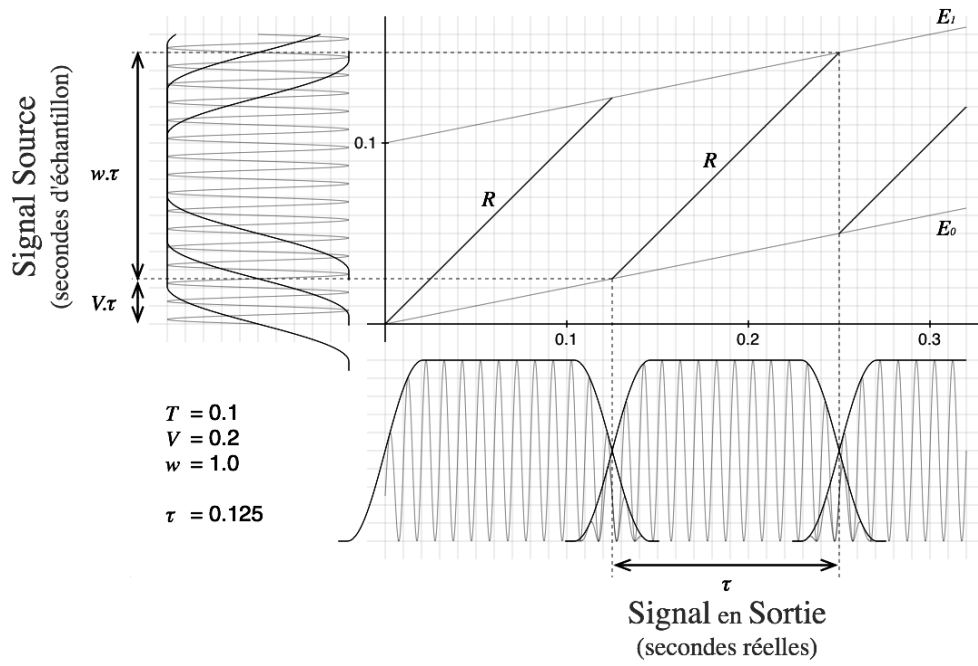


Figure 4 : Redistribution du Temps pour un Time-Stretch x 5

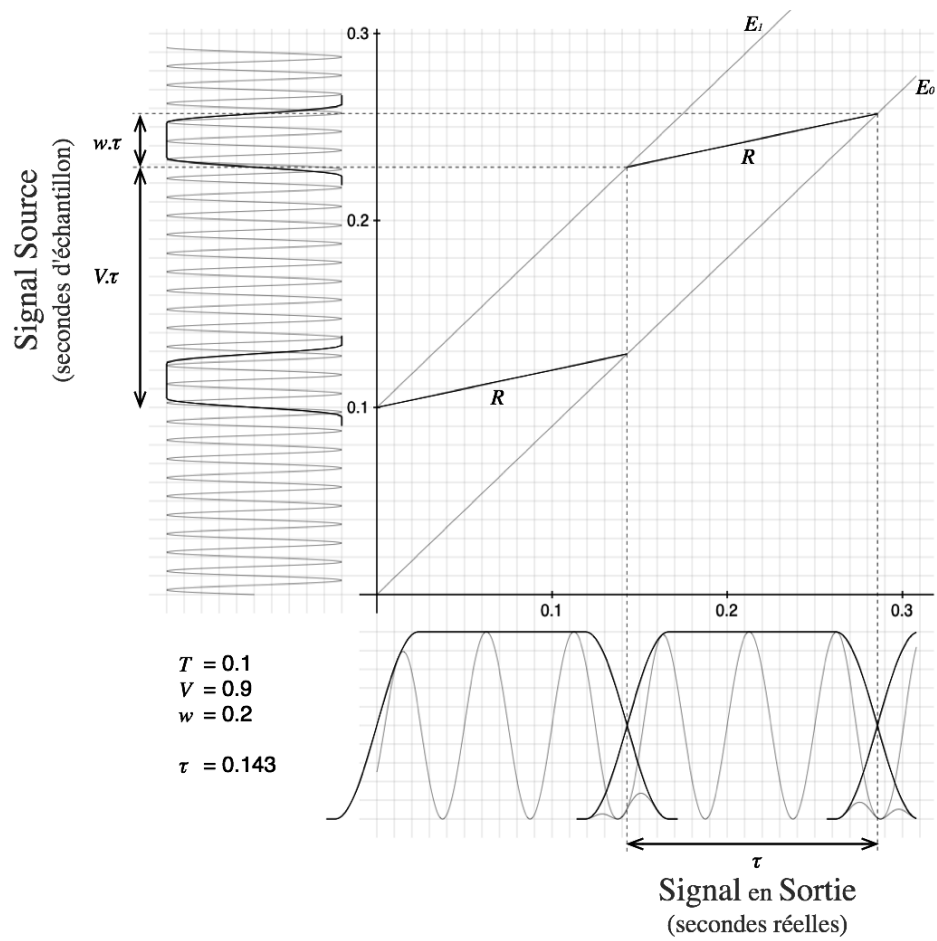


Figure 5 : Redistribution pour une Transposition / 5 et un peu de Dilatation

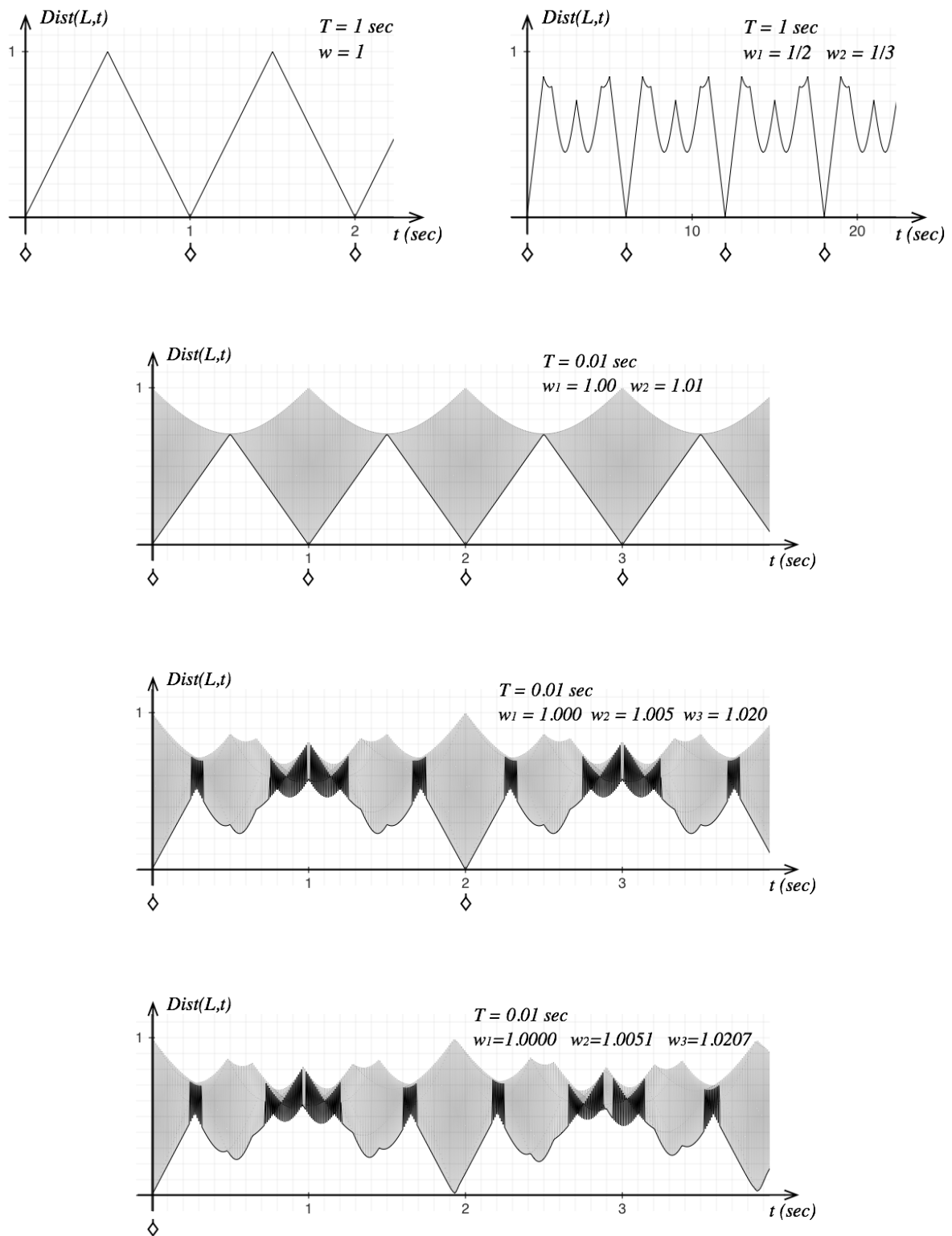


Figure 6 : Quelques Tracés de la Fonction $Dist$