Feuille d'exercices n^{o}10

Corrigé

Exercice 1

Pour tous k, l, en notant $\delta_n = 1$ si $n \equiv 0[2N]$ et $\delta_n = 0$ sinon, on a :

$$\begin{aligned}
\langle s_k, s_l \rangle &= \frac{\lambda_k \lambda_l}{2N} \sum_{n=0}^{N-1} \left(e^{ik\pi(n+1/2)/N} - e^{-ik\pi(n+1/2)/N}\right)\left(e^{-il\pi(n+1/2)/N} - e^{il\pi(n+1/2)/N}\right) \\
&= \frac{\lambda_k \lambda_l}{2N} \sum_{n=0}^{N-1} \left(e^{i(k-l)\pi(n+1/2)/N} - e^{-i(l+k)\pi(n+1/2)/N} - e^{i(k+l)\pi(n+1/2)/N} + e^{-i(k-l)\pi(n+1/2)/N}\right) \\
&= \frac{\lambda_k \lambda_l}{2N} \sum_{n=-N}^{N-1} \left(e^{i(k-l)\pi(n+1/2)/N} - e^{-i(l+k)\pi(n+1/2)/N}\right) \\
&= \lambda_k \lambda_l \left(\delta_{k-l} e^{i(k-l)\pi/(2N)} - \delta_{k+l} e^{-i(k+l)\pi/(2N)}\right)
\end{aligned}$$

On a toujours $\delta_{k-l} = \delta_{k+l} = 0$ si $k \neq l$ et $1 \leq k, l \leq N$.
Donc $\langle s_k, s_l \rangle = 0$ si $k \neq l$.
Si $k = l$, on a $\delta_{k-l} = 1$. On a $\delta_{k+l} = 1$ si et seulement si $k = l = N$. On en déduit que $||s_k||^2 = 1$.

Quand on décompose un signal dans la base des sinus, il y a presque toujours des coefficients "hautes-fréquences" (c'est-à-dire k assez grand) non-négligeables. En effet, les coefficients basses-fréquences sont petits en $n = 0$ et $n = N-1$. Si le signal auquel on s'intéresse ne s'annule pas au bord, sa décomposition sur la base de sinus fera intervenir des hautes fréquences.
C'est ennuyeux en traitement du signal car il est plus facile de compresser ou d'interpréter un signal qui est représenté par un petit nombre de coefficients.

Exercice 2

1. On prend la base $g_{m,m'}[n, n'] = g_m[n] g_{m'}[n']$. On a $\langle g_{m,m'}, g_{l,l'} \rangle = \langle g_m, g_l \rangle \langle g_{m'}, g_{l'} \rangle$. La base est donc toujours orthonormée.
Pour un signal $x[n, n']$ avec $0 \leq n, n' < N$, on a :

$$\langle x, g_{m,m'} \rangle = \sum_{n,n'} x[n, n'] g_m[n] g_{m'}[n'] = \sum_{n'} \langle x[., n'], g_m \rangle g_{m'}[n']$$

On calcule donc, pour tout n', la décomposition de $x[., n']$ sur la base g (c'est-à-dire qu'on calcule l'ensemble des $a[n', m] := \langle x[., n'], g_m \rangle$). Puisqu'il y a N valeurs pour n', cela nécessite $NC(N)$ opérations.
Ensuite, les $\langle x, g_{m,m'} \rangle$ s'obtiennent comme les coefficients des $a[., m]$ sur la base g. On a donc encore $NC(N)$ opérations.

Il y a donc au total $2NC(N)$ opérations.
2. La base de Fourier à deux dimensions est obtenue de cette manière et la transformée de Fourier rapide à deux dimensions se calcule de manière séparable à partir de la transformée de Fourier rapide à une dimension.

Exercice 3

1. Le cours dit que, si D_n est l'erreur quadratique engendrée par la quantification de A_n, la quantification optimale (D_n restant fixée) est la quantification uniforme, qui minimise l'entropie avec $H_n = H_d[A_n] - \frac{1}{2}\log_2(12D_n)$.
L'entropie de $(A_1, ..., A_N)$ (de manière équivalente, le nombre de bits nécessaire pour le codage) est donc :

$$H = \sum_n \left(H_d[A_n] - \frac{1}{2}\log_2(12D_n) \right)$$

Un autre théorème du cours dit que, la somme des D_n (c'est-à-dire la distorsion totale) étant fixée, le choix qui minimise l'entropie est de prendre tous les D_n égaux à D/N (où D est la distorsion totale). On a alors :

$$H = \sum_n H_d[A_n] - \frac{N}{2}\log_2(12D/N) = N\overline{H}_d - \frac{N}{2}\log_2(12D/N)$$

Lorsque D est fixée, choisir la base de façon à ce que l'entropie soit minimale revient donc à minimiser $\overline{H}_d$.
2. a) C'est $\log_2(\sqrt{2\pi e}) + \log_2(\sigma)$ (voir la question 1.b) de l'exercice 5).
b) $\overline{H}_d = \log_2(\sqrt{2\pi e}) + \frac{1}{N}\sum\limits_n \log_2(\sigma_n)$
c) On vérifie que $\sigma_n^2 = \langle g_n, Kg_n\rangle$: si on note $(u_1, ..., u_N)$ les coordonnées de g_n dans la base canonique $\sigma_n^2 = E(\langle g_n, X\rangle^2) = E((\sum\limits_k X_k u_k)^2) = \sum\limits_{k,l} u_k u_l E(X_k X_l) = \sum\limits_{k,l} u_k u_l K_{k,l} = \langle g_n, Kg_n\rangle$.
D'après la question b), il faut minimiser $\sum\limits_n \log_2(\sigma_n) = \frac{1}{2}\sum\limits_n \log_2(\sigma_n^2)$, ce qui revient à minimiser $\sum\limits_n \log_2(\langle g_n, Kg_n\rangle)$. Puisque $\log_2$ est strictement concave, le résultat découle de la question 3.
3. Soit $(e_1, ..., e_N)$ une base orthonormée de vecteurs propres de K (elle existe car K est symétrique et positive). On note λ_k les valeurs propres associées.
Pour tout n, on écrit $g_n = \sum\limits_k \alpha_k^n e_k$.
La matrice $(\alpha_k^n)_{1\le k,n\le N}$ est orthogonale : c'est une matrice de changement de base entre deux bases orthonormées.
Alors $\sum\limits_{n\le N} \phi(\langle g_n, Kg_n\rangle) = \sum\limits_n \phi\left(\sum\limits_k \alpha_k^{n2}\lambda_k\right)$. Puisque $\sum\limits_k \alpha_k^{n2} = 1$ (car les g_n sont de norme 1) et puisque ϕ est concave :

$$\begin{aligned}\sum_{n\le N} \phi(\langle g_n, Kg_n\rangle) &\ge \sum_{n,k} \alpha_k^{n2}\phi(\lambda_k) \\ &= \sum_k \phi(\lambda_k)\end{aligned}$$

(car $\sum\limits_n \alpha_k^{n2} = 1$ pour tout k)

L'égalité est atteinte si et seulement si on a l'égalité dans l'inégalité de concavité, c'est-à-dire qu'il faut, pour tout n, que tous les λ_k tels que $\alpha_k^n \neq 0$ soient identiques, ce qui revient à dire que g_n est un vecteur propre pour la valeur propre λ_k.

Exercice 4

1. Soit G telle que G est affine sur $[y_{k-1}; y_k]$, pour tout k, et envoie y_{k-1} sur $\tilde{y}_{k-1}$ et y_k sur $\tilde{y}_k$. Pour tout $k = 2, ..., K-1$, on a $G(a_k) = G\left(\frac{y_{k-1}+y_k}{2}\right) = \frac{G(y_{k-1})+G(y_k)}{2} = \frac{\tilde{y}_{k-1}+\tilde{y}_k}{2} = \tilde{a}_k$.
Sur $]-\infty; y_1]$, on choisit G affine de sorte que $G(y_1) = \tilde{y}_1$ et $G(a_1) = \tilde{a}_1$. De même sur $[y_K; +\infty[$.
Pour tout $t \in \mathbb{R}$, si on note k l'entier tel que $t \in [y_{k-1}; y_k]$, alors $G(t) \in [\tilde{y}_{k-1}; \tilde{y}_k]$ donc $\tilde{Q}(G(t)) = \tilde{a}_k$ et on a $G^{-1} \circ \tilde{Q} \circ G(t) = G^{-1}(\tilde{a}_k) = a_k = Q(t)$.
Sur l'intervalle $[y_{k-1}; y_k]$, on a $G'(t) = \frac{\tilde{y}_k - \tilde{y}_{k-1}}{y_k - y_{k-1}}$. C'est en particulier vrai pour $t = a_k$.

2. L'hypothèse que Q et $\tilde{Q}$ sont des quantificateurs haute résolution fait qu'on peut supposer que p est nulle en-dehors de $[y_1; y_{K-1}]$ (et constante sur $[y_{k-1}; y_k]$ pour tout k).

$$\begin{aligned}
D &= E(|X - Q(X)|^2) \\
&= \sum_{k=2}^{K-1} \int_{y_{k-1}}^{y_k} (x - a_k)^2 p(a_k) dx \\
&= \sum_{k=2}^{K-1} \frac{(y_k - y_{k-1})^3}{12} p(a_k) \\
&= \frac{1}{12} \sum_{k=2}^{K-1} \frac{1}{G'(a_k)^3} (\tilde{y}_k - \tilde{y}_{k-1})^3 p(a_k) \\
&= \frac{(\tilde{y}_{K-1} - \tilde{y}_1)^3}{12(K-2)^3} \sum_{k=2}^{K-1} \frac{1}{G'(a_k)^3} p(a_k)
\end{aligned}$$

On a utilisé le fait que $\tilde{y}_k - \tilde{y}_{k-1} = \frac{\tilde{y}_{K-1} - \tilde{y}_1}{K-2}$ car la subdivision $(\tilde{y}_k)_k$ est uniforme.

3.

$$\begin{aligned}
\sum_{k=2}^{K-1} \frac{1}{G'(a_k)^3} p(a_k) &= \sum_{k=2}^{K-1} \frac{1}{\tilde{y}_k - \tilde{y}_{k-1}} \frac{1}{G'(a_k)^2} p(a_k)(y_k - y_{k-1}) \\
&\approx \frac{K-2}{\tilde{y}_{K-1} - \tilde{y}_1} \int_{y_1}^{y_{K-1}} \frac{p(x)}{G'(x)^2} dx
\end{aligned}$$

donc :

$$D \approx \frac{(\tilde{y}_{K-1} - \tilde{y}_1)^2}{12(K-2)^2} \int_{y_1}^{y_{K-1}} \frac{p(x)}{G'(x)^2} dx$$

4. a) Par Hölder, $\int_{y_1}^{y_{K-1}} p(x)^{1/3} dx = \int_{y_1}^{y_{K-1}} \left(\frac{p(x)}{G'(x)^2}\right)^{1/3} (G'(x))^{2/3} dx \leq \left(\int \frac{p(x)}{G'(x)^2} dx\right)^{1/3} \left(\int G'(x) dx\right)^{2/3}$.
Donc $\int_{y_1}^{y_{K-1}} \frac{p(x)}{G'(x)^2} dx \geq \left(\int_{y_1}^{y_{K-1}} p(x)^{1/3} dx\right)^3 \left(\int_{y_1}^{y_{K-1}} G'(x) dx\right)^{-2} = \frac{\left(\int_{y_1}^{y_{K-1}} p(x)^{1/3} dx\right)^3}{(\tilde{y}_{K-1} - \tilde{y}_1)^2}$.
Ainsi :

$$D \geq \frac{1}{12(K-2)^2} \left(\int_{y_1}^{y_{K-1}} p(x)^{1/3} dx\right)^3$$

L'égalité est atteinte si G' est proportionnelle à $p^{1/3}$ (cas d'égalité de l'inégalité de Hölder).
b) Pour $\tilde{Q}$, on a $D \approx \frac{(\tilde{y}_{K-1}-\tilde{y}_1)^2}{12(K-2)^2}$ (cela se déduit de la question 3., pour $G' = 1$), ce qui peut être nettement plus grand si p n'est pas à peu près constante sur $[\tilde{y}_1; \tilde{y}_{K-1}]$.

Exercice 5

1. a) L'expression de D étant donnée dans l'énoncé, il suffit de calculer $\left(\int p(x)^{1/3}dx\right)^3$.
On a $p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)}$.

$$\begin{aligned}
\int p(x)^{1/3}dx &= \frac{1}{\left(\sqrt{2\pi}\sigma\right)^{1/3}} \int e^{-x^2/(6\sigma^2)}dx \\
&= \frac{1}{\left(\sqrt{2\pi}\sigma\right)^{1/3}} \sqrt{6\pi}\sigma \\
&= \sqrt{3}\left(\sqrt{2\pi}\sigma\right)^{2/3}
\end{aligned}$$

Donc $D \sim \frac{3^{3/2} 2\pi\sigma^2}{12K^2} = \frac{\sqrt{3}\pi}{2}\sigma^2 2^{-2R}$.
On a alors $C = \frac{\sqrt{3}\pi}{2}$.
b) On sait que le nombre de bits moyen optimal est à peu près l'entropie de la variable aléatoire quantifiée, soit (en reprenant les notations de l'exercice précédent) :

$$R_{opt} = -\sum_{k=2}^{K-1} P(a_k)\log_2(P(a_k))$$

avec $P(a_k) = p(a_k)(y_k - y_{k-1})$.
D'après la question 4.a) de l'exercice précédent, la subdivision optimale est obtenue lorsque G' est proportionnelle à $p^{1/3}$, c'est-à-dire lorsque $G' = \alpha p^{1/3}$ pour un certain réel α.
Le réel α est tel que :

$$\int_{y_1}^{y_{K-1}} G'(x)dx = \tilde{y}_{K-1} - \tilde{y}_1$$

c'est-à-dire :

$$\alpha = \frac{\tilde{y}_{K-1} - \tilde{y}_1}{\int_{\mathbb{R}} p^{1/3}(x)dx}$$

Alors :

$$\begin{aligned}
R_{opt} &= -\sum_{k=2}^{K-1} P(a_k)\log_2(p(a_k)(y_k - y_{k-1})) \\
&= -\sum_{k=2}^{K-1} P(a_k)\log_2\left(p(a_k)\frac{\tilde{y}_k - \tilde{y}_{k-1}}{G'(a_k)}\right) \\
&= -\sum_{k=2}^{K-1} P(a_k)\log_2\left(p(a_k)\frac{\tilde{y}_k - \tilde{y}_{k-1}}{\alpha p^{1/3}(a_k)}\right) \\
&= -\sum_{k=2}^{K-1} P(a_k)\log_2\left(p^{2/3}(a_k)\frac{\tilde{y}_k - \tilde{y}_{k-1}}{\tilde{y}_{K-1} - \tilde{y}_1}\int_{\mathbb{R}} p^{1/3}(x)dx\right)
\end{aligned}$$

$$
\begin{aligned}
&= -\sum_{k=2}^{K-1} P(a_k)\log_2\left(\frac{p^{2/3}(a_k)}{K-2}\int_{\mathbb{R}} p^{1/3}(x)dx\right)\\
&= \sum_{k=2}^{K-1} P(a_k)\log_2(K-2) - \sum_{k=2}^{K-1} P(a_k)\log_2\left(p^{2/3}(a_k)\int_{\mathbb{R}} p^{1/3}(x)dx\right)\\
&= \log_2(K-2) - \sum_{k=2}^{K-1} p(a_k)\log_2\left(p^{2/3}(a_k)\int_{\mathbb{R}} p^{1/3}(x)dx\right)(y_k - y_{k-1})\\
&= \log_2(K-2) - \int_{\mathbb{R}} p(x)\log_2\left(p^{2/3}(x)\int_{\mathbb{R}} p^{1/3}(x)dx\right)
\end{aligned}
$$

donc, lorsque K augmente, R_{opt} est égal à $\log_2(K)$, à une constante près. On ne peut donc pas utiliser significativement moins de bits.
c) Pour un quantificateur uniforme, le cours dit qu'on a :

$$D = \frac{\Delta^2}{12}$$

où Δ est le pas de subdivision.
D'autre part, toujours d'après le cours, l'entropie (donc le nombre de bits moyen utilisé pour le codage) vaut, lorsque le quantificateur est uniforme :

$$R = H_d - \frac{1}{2}\log_2(12D) = H_d - \log_2(\Delta)$$

Il faut donc calculer l'entropie différentielle H_d :

$$
\begin{aligned}
H_d &= \int p(x)\log(p(x))dx\\
&= -\int \frac{e^{-x^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma}\log_2\left(\frac{e^{-x^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma}\right)dx\\
&= -\int \frac{e^{-x^2/2}}{\sqrt{2\pi}}\log_2\left(\frac{e^{-x^2/2}}{\sqrt{2\pi}\sigma}\right)dx\\
&= -\int \frac{e^{-x^2/2}}{\sqrt{2\pi}}\log_2\left(\frac{e^{-x^2/2}}{\sqrt{2\pi}}\right)dx + \int \frac{e^{-x^2/2}}{\sqrt{2\pi}}\log_2(\sigma)dx\\
&= \log_2(\sqrt{2\pi e}) + \log_2(\sigma)
\end{aligned}
$$

On a donc :

$$
\begin{aligned}
R &= \log_2(\sqrt{2\pi e}) + \log_2(\sigma) - \log_2(\Delta)\\
\implies \Delta &= 2^{\log_2(\sqrt{2\pi e})+\log_2(\sigma)-R} = \sqrt{2\pi e}\sigma 2^{-R}
\end{aligned}
$$

On a alors :

$$D \sim \frac{\pi e}{6}\sigma^2 2^{-2R}$$

Pour un taux de codage de R, la distorsion quadratique est donc aussi en $C'\sigma^2 2^{-2R}$, mais on peut vérifier que $C' \approx 1,4$ alors que $C \approx 2,7$. Le codage uniforme est donc meilleur (le cours

indiquait qu'il était optimal) ; il permet, à nombre moyen de bits égal, de réduire la distorsion quadratique d'un facteur constant.
2. a) Pour un R_i donné, la quantification qui va minimiser la distorsion D_i sera la quantification uniforme. D'après la question 1.b), on aura :

$$D_i \approx \frac{\pi e}{6} \sigma_i^2 2^{-2R_i}$$

Il faut donc minimiser $\sum\limits_i \frac{\pi e}{6} \sigma_i^2 2^{-2R_i}$, sous la contrainte que $R_1 + ... + R_N = N\overline{R}$.
Il faut donc minimiser $\sum\limits_i \sigma_i^2 2^{-2R_i}$. Par l'inégalité arithmético-géométrique :

$$\sum_i \sigma_i^2 2^{-2R_i} \geq N \left(\prod_i \sigma_i^2 2^{-2R_i} \right)^{1/N} = N \left(\prod_i \sigma_i^2 \right)^{1/N} 2^{-2\overline{R}}$$

L'égalité est atteinte lorsque tous les $\sigma_i^2 2^{-2R_i}$ sont égaux (c'est-à-dire lorsque le pas de subdivision est le même selon chaque coordonnée).
La distorsion minimale vaut donc :

$$D \sim \frac{\pi e N}{6} \left(\prod_i \sigma_i^2 \right)^{1/N} 2^{-2\overline{R}}$$

b) Tous les R_i sont égaux à $\overline{R}$. On a alors :

$$\begin{aligned} D &= \sum_i D_i \\ &= \frac{\pi e}{6} \sum_i \sigma_i^2 2^{-2R_i} \\ &= \frac{\pi e}{6} \left(\sum_i \sigma_i^2 \right) 2^{-2\overline{R}} \end{aligned}$$

C'est donc moins bon que le quantificateur de la question précédence, car $\left(\prod\limits_i \sigma_i^2 \right)^{1/N} \leq \frac{1}{N} \sum\limits_i \sigma_i^2$.
La différence entre les deux quantificateurs est surtout sensible lorsqu'il y a de grandes variations entre les σ_i.