

RAPPORT 953-2*

**CODAGE NUMÉRIQUE POUR L'ÉMISSION DE SIGNAUX SON DE HAUTE QUALITÉ
EN RADIODIFFUSION PAR SATELLITE
(BANDE PASSANTE NOMINALE DE 15 kHz)**

(Question 2/10 et 11, Programmes d'études 51C/10 et 2F/10 et 11)

(1982-1986-1990)

1. Introduction

L'évolution de la technique permet actuellement d'envisager l'émission vers le public de signaux audio codés numériquement vers le public et il sera notamment possible de le faire dans le service de radiodiffusion par satellite. Cela demande la définition préalable des caractéristiques de codage, de multiplexage et de modulation des signaux audio. Le présent Rapport ne concerne que le codage audionumérique; le multiplexage, les méthodes de protection contre les erreurs sur les bits et la modulation sont traités dans les Rapports 954 et 632. Des informations concernant les méthodes de correction et de masquage des erreurs peuvent être obtenues dans le Rapport 1073.

Le Programme d'études 51C/10 souligne que l'introduction des techniques numériques pour l'émission du son doit permettre d'améliorer la qualité des signaux transmis. La Recommandation 651 traite du codage numérique MIC. Des résultats d'essais objectifs et subjectifs sont décrits dans l'Annexe II du présent Rapport.

D'autres méthodes de codage permettant une diminution importante du débit binaire sont décrites dans le Rapport 1199.

En plus de l'amélioration de la qualité, les considérations suivantes doivent être prises en compte:

- les compromis entre les objectifs de qualité et le débit binaire peuvent être différents pour les services sonores qui peuvent avoir diverses exigences de qualité et diverses contraintes de planification; des compromis peuvent également varier selon les besoins de chaque pays;
- l'utilisation d'une norme unique pour chaque application présente des avantages évidents pour les radiodiffuseurs, les constructeurs de récepteurs et le public.

2. Codage numérique MIC**2.1 Codage linéaire**

Le signal audio est codé sous forme numérique par un convertisseur analogique-numérique de haute précision après passage dans un filtre approprié anti-repliement, et éventuellement dans un réseau de préaccentuation. Le codage uniforme ainsi obtenu permet d'aboutir à une représentation linéaire avec un minimum de 14 bits par échantillon. Dans ce cas, on utilise un codage en complément à deux pour chaque échantillon. On peut également associer un facteur d'échelle à des blocs d'échantillons successifs.

2.2 Virgule flottante ou codage quasi instantané

Un signal son codé linéairement sous 16 bits peut être transmis à raison de 14 bits par échantillon au moyen d'un système de codage à virgule flottante dans lequel le facteur d'échelle est obtenu sur la base de 64 échantillons consécutifs. On peut utiliser ou non la préaccentuation.

Lorsqu'il est nécessaire pour le canal audio de recourir à un débit binaire inférieur par canal audio, la compression-extension quasi instantanée permet de réduire de 14 à 10 le nombre de bits par échantillon. Cette compression-extension est appliquée à des blocs de 32 échantillons successifs avec un facteur d'échelle complémentaire à 5 niveaux. On continue d'utiliser une représentation en complément de deux pour chaque échantillon. Dans ce cas, une préaccentuation est recommandée si l'on veut réduire le bruit de modulation.

* Ce Rapport doit être porté à l'attention de la CMTT.



2.3 Codage par transition

Le codage par transition est une autre technique pour obtenir un débit binaire réduit par voie audiofréquence; il repose sur le fait que dans la plupart des programmes radiophoniques, la corrélation entre deux échantillons consécutifs est élevée. En pareil cas, la différence entre échantillons successifs est en général moindre que les valeurs réelles des échantillons. En codant le signal différentiel au lieu des échantillons originaux, on peut réduire le débit binaire en fonction du rapport signal/bruit nécessaire.

Des échantillons différentiels peuvent être codés instantanément un par un, comme dans un codeur linéaire, ou par blocs d'échantillons successifs (avec un facteur d'échelle complémentaire) comme dans la compression-extension quasi instantanée. Cette combinaison du codage par transition et de la compression-extension quasi instantanée a été proposée [CCIR, 1986-90] afin d'obtenir une réduction (de 15 à 8) du nombre de bits par échantillon. L'utilisation de la préaccentuation est recommandée pour réduire le bruit de modulation.

2.4 *Accentuation MIC*

La qualité du système de transmission, notamment en ce qui concerne les caractéristiques de bruit et de distorsion, dépend largement des statistiques relatives aux signaux et de la loi de codage utilisée. En outre, on peut considérer que la sensibilité du système de réception (y compris l'oreille humaine) au bruit audible n'est pas uniforme. Compte tenu de ces facteurs, il serait possible d'améliorer les caractéristiques en utilisant à la fois une préaccentuation appropriée avant le processus de codage et la désaccentuation correspondante au récepteur après le processus de décodage. Dans le cas des systèmes MIC, deux méthodes d'accentuation de ce genre ont été étudiées et évaluées exhaustivement. L'Annexe I donne des renseignements plus détaillés au sujet de deux systèmes, dont l'un est fondé sur la Recommandation J.17 du CCITT et l'autre connu sous le nom de 50/15 μ s.

3. **Modulation delta adaptative (MDA)**

3.1 *Description générale*

Ce système est fondé sur la modulation delta et désigné comme système MDA (la définition mathématique du codeur est donnée au § 3.2). Le signal audiofréquence est codé sous forme numérique au moyen d'un modulateur delta simple après avoir été modifié par deux traitements:

- le signal audiofréquence traverse un réseau de préaccentuation variable qui altère son spectre en fréquence, et
- ensuite, il est comprimé en niveau selon sa pente. La compression est «infinie», dans une plage de 48 dB, c'est-à-dire que le signal d'entrée se trouve comprimé au même niveau pour être mis sous forme numérique. Une charge optimale de la voie numérique est obtenue pour la plupart des signaux audiofréquence.

Le train binaire représentant le signal audiofréquence codé est ensuite acheminé vers le récepteur, en même temps que deux trains numériques à très faible débit contenant des signaux de commande afin que le décodeur puisse effectuer avec précision le traitement inverse de celui du codeur. Le décodeur rétablit le signal audiofréquence en procédant à l'intégration du train binaire audiofréquence, et en effectuant une extension de la gamme dynamique au moyen du signal de commande de compression (pente), et procède à une désaccentuation spectrale à partir des informations contenues dans le signal de commande d'accentuation.

On peut noter les aspects ci-après en ce qui concerne l'utilisation de la modulation delta:

- a) il est possible de diminuer le débit binaire de la modulation delta, avec une incidence mineure sur la qualité des signaux audio (le rapport signal/bruit se dégrade de 9 dB pour un débit binaire réduit de moitié);
- b) la reproduction d'erreurs est tolérable;

- c) le convertisseur numérique-analogique pour la modulation delta est un dispositif très simple ne nécessitant aucun composant de précision;
- d) étant donné que la désaccentuation permet d'obtenir un filtrage adéquat à la sortie, il n'est pas nécessaire d'utiliser des filtres passe-bas à coupure brusque.

Les éléments décrits aux points c) et d) ont une incidence sur les circuits de décodage qui doivent être utilisés.

3.2 Définition du décodeur MDA (Fig. 1)

3.2.1 Décodeur audiofréquence

Le décodeur audiofréquence comporte un intégrateur à fuites alimenté par des impulsions prélevées sur le train binaire contenant les informations audiofréquences. Les données arrivant à un débit de l'ordre de 250 kbit/s à la sortie de l'intégrateur donnent un signal qui évolue positivement ou négativement en échelon. L'amplitude des impulsions est linéairement proportionnelle au signal de commande appliqué, sur une gamme d'environ 50 dB, et l'intégrateur à fuites a une réponse en fréquence définie par l'expression suivante:

$$(1 + sT_0)^{-1}$$

où $T_0 = 0,5$ ms.

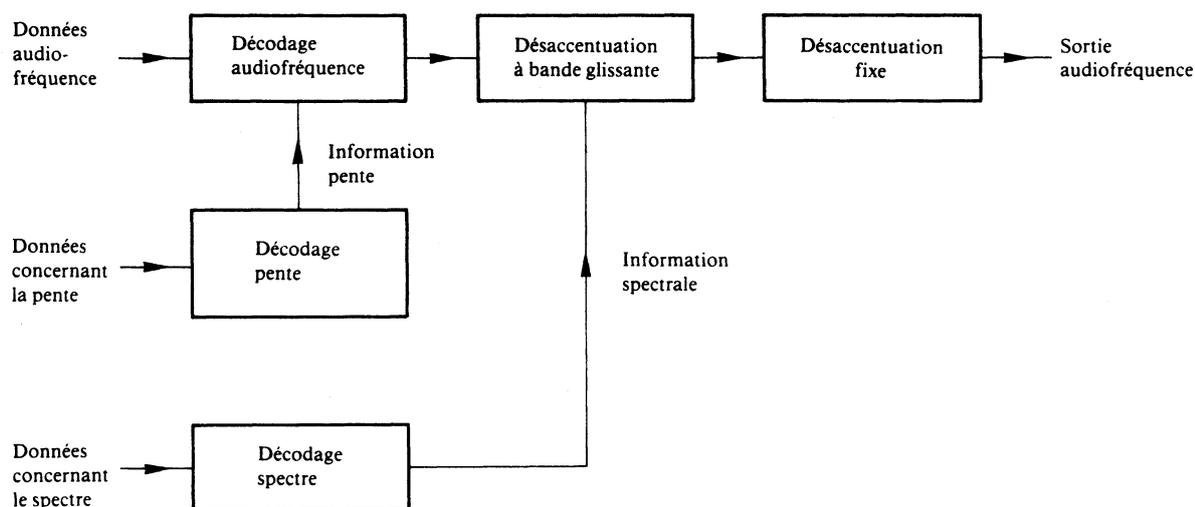


FIGURE 1 – Décodeur de modulation delta adaptative

3.2.2 Désaccentuation à bandes glissantes (Fig. 2)

Le système de désaccentuation a une réponse en fréquence variable qui est donnée par l'expression suivante:

$$\left[\frac{10sT_1}{1+sT_1} + \frac{1+sT_2}{1+sT_3} \right]^{-1}$$

où $T_2 = 5 \mu\text{s}$, $T_3 = 50 \mu\text{s}$ et T_1 est un temps variable commandé par les données spectrales.

Remarquez que le premier terme entre crochets est une caractéristique passe-haut à la fréquence f_1 où:

$$f_1 = \frac{1}{2nT_1}$$

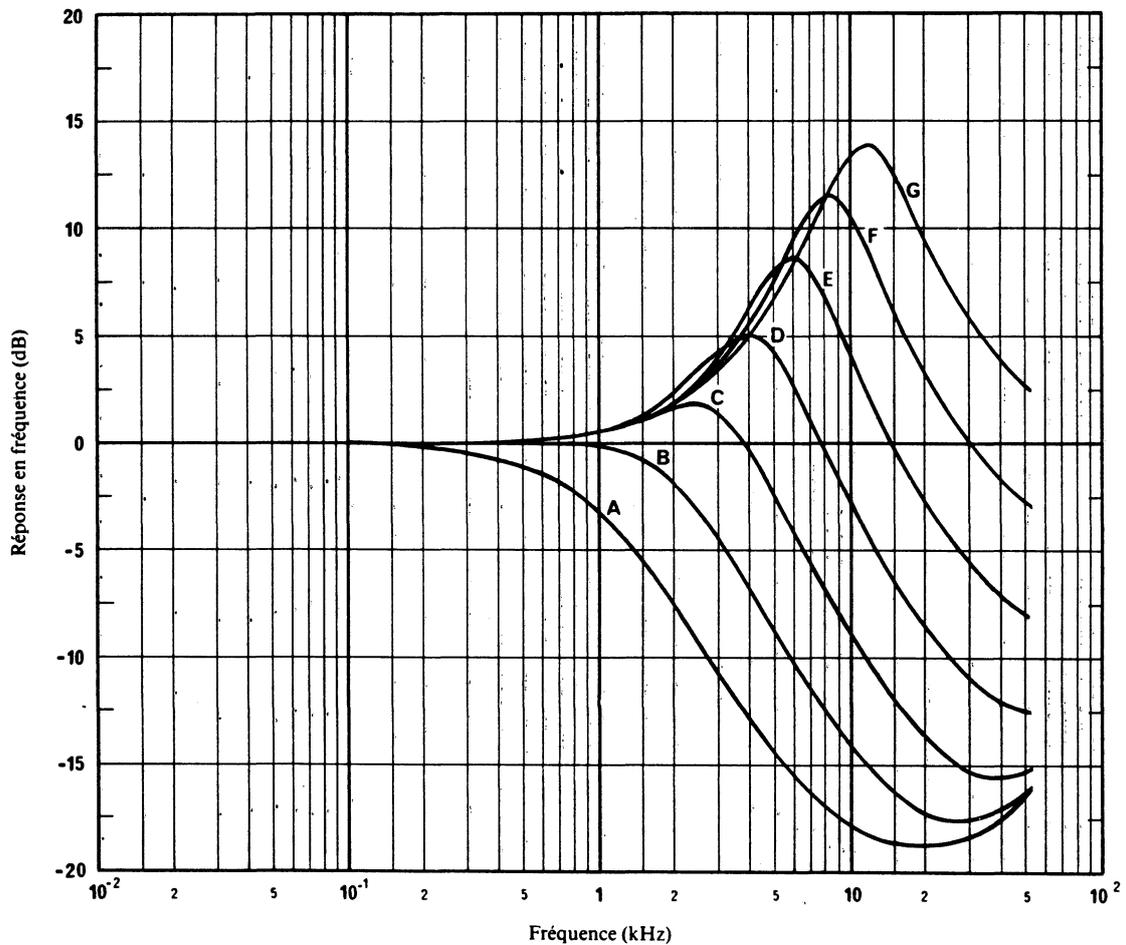


FIGURE 2 - *Filtre de désaccentuation MDA*

3.2.3 *Décodeur de spectre*

L'information de commande de la désaccentuation est contenue dans le train binaire à environ 8 kbit/s qui, en principe, mais pas nécessairement, est un sous-multiple entier du débit binaire du signal audiofréquence. Le niveau moyen V_m du train binaire est obtenu en acheminant les données au travers d'un filtre passe-bas à trois pôles dont la caractéristique est la suivante:

$$(1 + sT_4)^{-3}$$

où $T_4 = 2$ ms.

Si V_p est l'amplitude des impulsions de données, le paramètre de commande x est donné par l'expression:

$$x = \frac{V_m}{V_p}$$

Le signal à la sortie du filtre est ensuite dirigé vers un exponentiateur afin d'obtenir le signal de commande de l'accentuation variable. Les constantes du circuit sont telles qu'elles vérifient la relation:

$$f_1 = \frac{1}{2nT_1} = 4000(2^{10x}) \quad \text{Hz}$$

Avec cette définition, une modification de 0,1 de x déplace la fréquence f_1 d'une octave.

3.2.4 *Décodeur de pente*

L'information relative à la pente du signal se trouve dans un autre train de données à 8 kbit/s environ. Le train binaire est converti en un signal de commande au moyen d'un filtre passe-bas et d'un exponentiateur exactement comme dans le cas du décodeur de spectre. L'amplitude des impulsions qui sont intégrées dans le décodeur audiofréquence est directement proportionnelle au signal de commande.

Comme dans le cas du décodeur de spectre, le niveau moyen V_m du train binaire est obtenu en utilisant un filtre passe-bas à trois pôles dont la caractéristique est:

$$(1 + sT_4)^{-3}$$

où $T_4 = 2$ ms.

Le signal ainsi obtenu est ensuite dirigé sur un exponentiateur afin d'obtenir le signal de commande d'amplitude des impulsions de longueur de l'échelon V_{ss} destiné au décodeur audiofréquence. La caractéristique de cet exponentiateur est donnée par l'expression suivante:

$$V_{ss} = V_0(2^{10y})$$

où y désigne le niveau moyen normalisé des impulsions (comme x ci-dessus) et V_0 est un facteur de proportionnalité constant adapté au décodeur audiofréquence. Avec la présente définition, lorsque y varie de 0,1, l'amplitude de l'impulsion varie de 6 dB.

3.2.5 *Désaccentuation fixe*

Le circuit de désaccentuation fixe est constitué par un simple filtre passe-bas à un seul pôle présentant la caractéristique suivante:

$$(1 + sT_5)^{-1}$$

où $T_5 = 25$ μ s.

Avec les décodeurs grand public, aucun filtrage supplémentaire à la sortie n'est nécessaire.

REFERENCES BIBLIOGRAPHIQUES

Documents du CCIR:

[1986-90]: 10/52 (Japon).

BIBLIOGRAPHIE

- FIELDER, L. D. [septembre 1985] Pre- and post-emphasis techniques for audio recorders. *J. Audio Eng. Soc.*, Vol. 33, 9.
- FIELDER, L. D. [octobre 1985] Audible modulation noise in floating point conversion systems. *J. Audio Eng. Soc.*, Vol. 33, 10.
- FIELDER, L. D. [octobre 1985] Assessing the low level sound quality of digital audio systems. Présenté à la 19th Convention of the Audio Engineering Society, preprint No. 2268.
- TODD, C. et GUNDRY, K. [mars 1984] A digital audio system for broadcast and prerecorded media. Audio Engineering Society, preprint No. 2071, Paris.
- TODD, C. et GUNDRY, K. A digital audio system coding method for DBS, cable and terrestrial broadcasting. Conference Publication 240, Ninth International Broadcasting Convention (IBC '84), Brighton, Royaume-Uni.

ANNEXE I

TECHNIQUES D'ACCENTUATION APPLICABLES AUX SIGNAUX SON MIC DE HAUTE QUALITÉ

Dans la technique de codage MIC du son numérique, il faut utiliser des méthodes de pré- et de désaccentuation, à plusieurs fins:

- en cas de traitement du signal son numérique, l'accentuation diminue la perceptibilité subjective du bruit de quantification, surtout en ce qui concerne le bruit «modulé par le programme» dans le cas du codage avec compression-extension;
- dans le cas de traitement des signaux son avec compression-extension et avec codage linéaire, l'accentuation peut diminuer la dégradation due au taux d'erreur binaire pour des valeurs peu élevées du rapport C/N .

Deux caractéristiques différentes ont été proposées pour les systèmes de codage à compression-extension et linéaire. Le premier système de préaccentuation est fondé sur la Recommandation J.17 du CCITT (avec un affaiblissement d'insertion de 6,5 dB à 0,8 kHz) et le second système est fondé sur la constante de temps 50/15 μ s. Les caractéristiques de ces systèmes sont représentées à la Fig. 1 de la Recommandation 651.

Les résultats d'études concernant le codage [CCIR, 1982-86a] ainsi que le codage linéaire avec compression-extension [CCIR, 1982-86b et c] montrent clairement quels sont les avantages de la préaccentuation dans la plupart des cas. On a observé que les deux systèmes offrent l'avantage de réduire le bruit de quantification, et que la préaccentuation à 50/15 μ s est légèrement plus avantageuse dans le cas où la compression-extension n'est pas utilisée. S'agissant des risques de surcharge, les documents [CCIR, 1982-86a, b et c] exposent des points de vue différents. Aucune différence nette n'a pu être constatée entre les deux systèmes considérés pour les dégradations dues au taux d'erreur binaire quand les valeurs du rapport C/N sont peu élevées.

Selon des évaluations théoriques [CCIR, 1982-86b] et des essais subjectifs [CCIR, 1982-86d et e] effectués pour des systèmes à compression-extension fondés sur les caractéristiques de préaccentuation représentées à la Fig. 1 de la Recommandation 651, on pourrait obtenir une importante réduction de l'audibilité subjective du bruit modulé par le programme en utilisant la préaccentuation décrite dans la Recommandation J.17 du CCITT. Toutefois, les résultats obtenus à la suite d'autres essais subjectifs [CCIR, 1982-86f et g] n'ont fait apparaître aucune différence appréciable. Le document [CCIR, 1982-86f], précise qu'un ajustement de l'affaiblissement d'insertion a été fait pour chaque système afin d'éviter les surcharges.

En ce qui concerne le codage numérique des signaux de programmes sonores, on utilise largement les dispositions de la Recommandation J.17 du CCITT (avec un affaiblissement d'insertion de 6,5 dB à 0,8 kHz) aux fins de la transmission, comme il est spécifié dans la Recommandation J.41 du CCITT, et l'on utilise beaucoup le système de préaccentuation à 50/15 μ s pour les applications grand public.

RÉFÉRENCES BIBLIOGRAPHIQUES

Documents du CCIR

[1982-86]: a. 10/21 (UER); b. 10-11S/201 (UER); c. 10-11S/139 (Japon); d. 10/269 (Allemagne (République fédérale d')); e. 10-11S/206 (France); f. 10-11S/205 (Japon); g. 10-11S/207 (Canada).

ANNEXE II

Résultats d'essais subjectifs de la modulation delta
adaptative et d'autres méthodes employées
pour le codage son de haute qualité

1. Mesures subjectives réalisées en Australie

L'Australian Broadcasting Corporation (AuBC) a réalisé des essais subjectifs sur le système de modulation delta adaptative (MDA) [AuBC, 1985] conformément à la Recommandation 562, afin d'établir une relation entre la qualité subjective du son et le taux d'erreur binaire, avec comme objectif éventuel de relier ces résultats avec le rapport C/N .

L'équipement de modulation delta adaptative comportait un générateur aléatoire d'erreurs qui permettait d'introduire des erreurs dans le train de données avec un taux parfaitement connu et sélectionné au moyen d'un commutateur. La précision dans la sélection a été vérifiée et on a constaté qu'elle était suffisamment précise à l'intérieur des variations statistiques normales.

Pour reproduire les éléments sonores en séquences, conformément à la Recommandation 562, on a choisi des conditions soigneusement contrôlées et mesurées.

Pour l'évaluation de la partie son du système, sept séquences différentes ont été choisies sur disque compact ou dans des enregistrements originaux MIC à 16 bits, puis enregistrées successivement en MIC à 16 bits sur un magnéscope, pour des reproductions ultérieures. Les sept séquences différentes de programme sont données au Tableau 1.

TABLEAU I. – Séquences de programme ayant servi de référence

Numéro (courbe)	Description
1	Solo masculin
2	Mot prononcé par une voix masculine
3	Piano
4	Flûte puis jeunes choristes masculins
5	Mot prononcé par une voix féminine
6	Orchestre
7	Groupe moderne mixte

La Fig. 3 montre que pour un TEB donné, les dégradations varient considérablement mais régulièrement en fonction du contenu du programme; toutefois, un taux de dégradation moyen de 3,5 est obtenu pour un TEB d'environ 1×10^{-3} .

On a estimé que la note de dégradation générale de 3,5 pour un TEB de 1×10^{-3} était équivalente à un rapport C/N inférieur à 9 dB. Cependant, la note serait plus basse si on utilisait quatre niveaux de données, comme cela a été proposé pour l'Australie. Il n'a pas été possible d'effectuer des mesures sur ce mode de fonctionnement, mais les résultats obtenus, montrent une dégradation d'environ 1 dB pour les niveaux externes de données et d'environ 1,8 dB pour les niveaux internes. Ceci pourrait avoir des conséquences quant aux genres de programmes à diffuser dans telle ou telle voie de données. Toutefois, l'AuBC ne disposait pas des appareils qui auraient permis de corroborer ces chiffres.

Il convient d'établir une corrélation indépendante entre le rapport C/N et le TEB obtenu avec des récepteurs en service opérationnel.

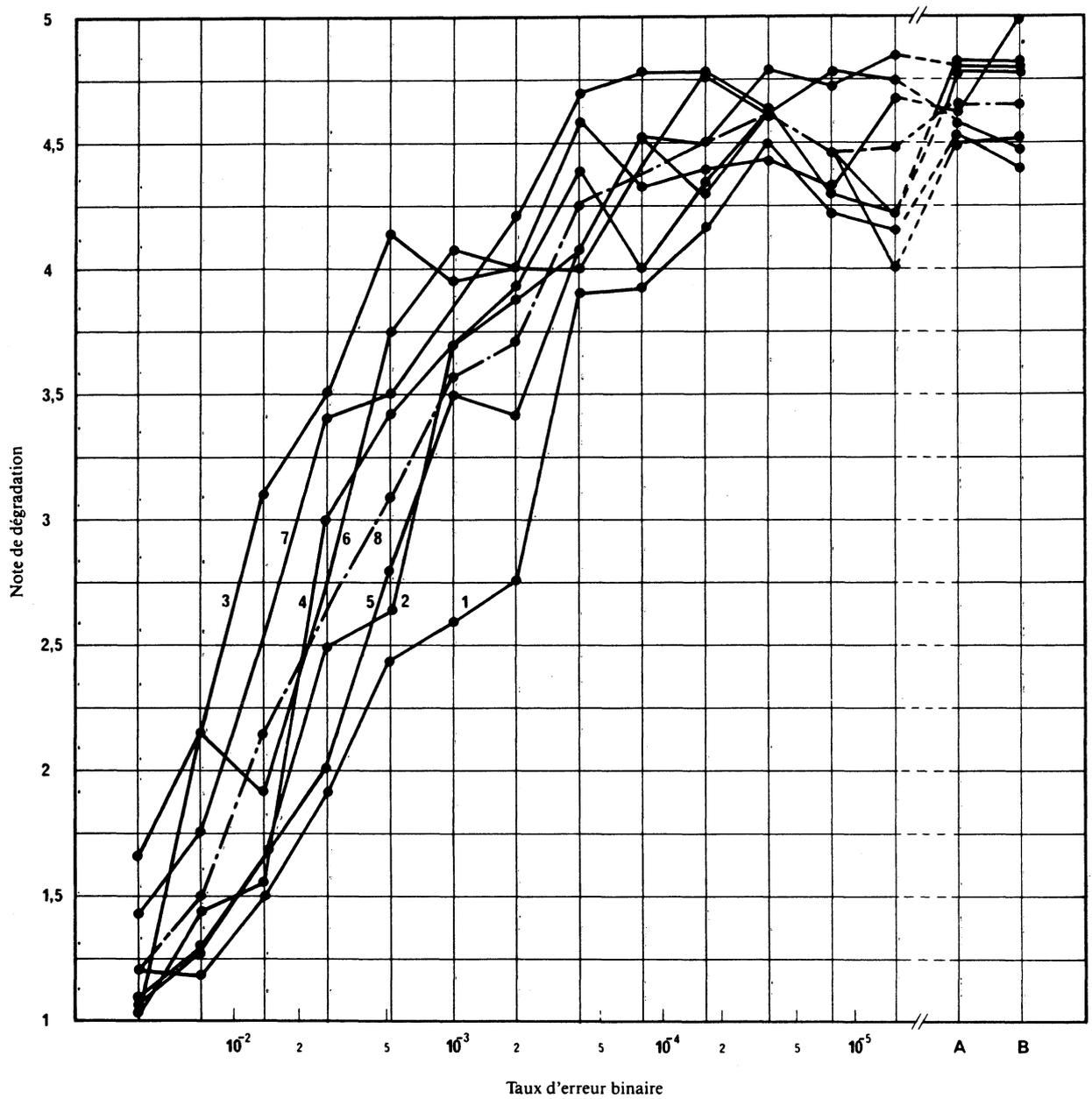


FIGURE 3 – Niveau de dégradation en fonction de la position du commutateur, rubrique par rubrique

Note 1. – Les courbes N^{os} 1-7 correspondent aux rubriques N^{os} 1-7 énumérées dans le Tableau III: la Courbe 8 correspond au résultat global (moyen).

Note 2. – Dans la position A du commutateur d'échelle, le signal est émis par le système Dolby ADM sans erreurs.

Note 3. – Dans la position B du commutateur d'échelle, le signal vient directement de la source analogique (ex. MIC sur magnéto-scope à cassette).

2. Mesures subjectives effectuées au Canada

2.1 Evaluation subjective de divers codeurs MDA, linéaires et quasi instantanés

Un programme d'essais a été réalisé au Canada dans le but de déterminer les avantages relatifs des différentes méthodes de codage proposées.

Au cours de ce programme, les méthodes de codage suivantes ont été étudiées:

- Compression-extension quasi instantanée 10-14 (NICAM 3) avec la préaccentuation donnée dans la Recommandation J.17 du CCITT.
- Compression-extension instantanée 10-14 (loi A) avec la préaccentuation donnée dans la Recommandation J.17 du CCITT.
- Codage linéaire à 14 bits avec la préaccentuation donnée dans la Recommandation J.17 du CCITT.
- Modulation delta adaptative (fréquence d'échantillonnage: 330 kHz).
- Modulation delta adaptative (fréquence d'échantillonnage: 204 kHz).

Les trois premières méthodes de codage ont été modélisées sur un ordinateur couplé à un processeur vectoriel.

La préaccentuation appliquée au signal a été modifiée au moyen d'un filtre numérique. La fréquence d'échantillonnage a également été modifiée à 32 kHz par traitement numérique. Une compression-extension suivie d'écrêtage permettant de se ramener au nombre de bits approprié a ensuite été effectuée afin d'obtenir le train numérique. La remise au format approprié a été effectuée afin d'enregistrer la séquence de test sur un magnétoscope.

Ces traitements n'ont pas été utilisés pour le système ADM qui n'a pu être modélisé sur l'ordinateur. Il était donc nécessaire de convertir le signal sous une forme analogique avant qu'il ne soit traité par le dispositif de codage ADM. A la sortie de l'ADM, le signal était à nouveau converti de sa forme analogique en information numérique à 16 bits pour être inséré dans la séquence d'essai enregistrée.

Les séquences d'essai étaient, soit lues à partir de disques compacts, soit produites par l'ordinateur lui-même:

- Cardozo, Pájaro Campana (Harpe indienne)
- Orff, Carmina Burana (Chant)
- Gong synthétique produit à 44 kHz simulant le son d'un triangle sur les 4 premières notes de «Frère Jacques» (4,2 kHz, 4,7 kHz, 5,3 kHz et 4,2 kHz). (Temps d'attaque: $T = 3$ ms, amortissement: $T = 300$ ms)
- Gong synthétique produit à 44 kHz simulant une guitare basse sur les 4 premières notes de «Frère Jacques» (65,4 Hz, 73,4 Hz, 82,4 Hz et 65,4 Hz). (Temps d'attaque: $T = 3$ ms, amortissement: $T = 300$ ms)

Les essais ont été effectués selon la méthode de comparaison proposée par le CCIR dans la Recommandation 562 où la séquence A-B est présentée deux fois et où l'on laisse 15 s pour donner une note. On a également utilisé l'échelle de notation du CCIR. Les mesures audiométriques ont été effectuées sur chaque sujet avant les essais subjectifs. En moyenne, la salle d'écoute qui était bien étalonnée pouvait recevoir 3 auditeurs. Les essais ont été réalisés avec 25 auditeurs au total.

Toutes les comparaisons ont été répétées en ordre inverse afin d'établir la cohérence des auditeurs. Ainsi, 4 auditeurs ont été écartés pour l'analyse des résultats. De plus, lorsque les deux ensembles de résultats d'un même auditeur pour une même comparaison présentaient une différence de plus de deux points, ces derniers n'étaient pas pris en considération.

En raison des vibrations et des tremblements provoqués dans la salle d'écoute par le niveau élevé du son du gong synthétisé à basse fréquence, la séquence d'essai a été présentée à 8 auditeurs en utilisant des écouteurs. Le résultat de ce groupe concerne le gong à basse fréquence.

En plus des résultats moyens d'opinion et de l'écart type obtenu à partir des données filtrées, une étude sur la signification statistique des résultats a été menée en utilisant une distribution «Student-t à une queue». Le niveau de confiance, en moyenne, montre qu'une population normale préfère le codec «A» au codec «B»*.

Le Tableau IF résume les résultats. L'opinion moyenne exprimée sur l'échelle -3, 0, +3, l'écart type ainsi que le niveau de confiance indiquant que le «codec A est meilleur que le codec B» sont donnés pour chacun des cas. Dans le tableau, la compression-extension quasi instantanée 10-14 décrite dans le Rapport 953, est désignée en tant que NICAM 3.

* «codec» = «codeur-décodeur».

TABLEAU I I – Résultats des essais subjectifs effectués au Canada

		Harpe	Vocal	Gong HF	Gong BF
Codec «A»	Codec «B»	$\bar{X} / S / \% \text{ Conf}$			
NICAM 3, Rec. J.17 du CCITT	ADM (330 kHz)	0,1 0,7 88%	0,2 0,9 91%	-0,1 1,3 28%	1,1 0,8 99,95%
NICAM 3, Rec. J.17 du CCITT	ADM (204 kHz)	0,0 0,8 50%	0,2 0,9 92%	1,8 1,0 99,999%	1,4 0,5 99,995%
ADM (330 kHz)	Loi A, Rec. J.17 du CCITT	0,0 0,6 50%	-0,2 0,8 6%	0,9 1,4 99,999%	-1,2 0,5 0,01%
NICAM 3, Rec. J.17 du CCITT	Loi A, Rec. J.17 du CCITT	0,0 0,7 50%	-0,1 1,0 27%	0,8 1,1 99,975%	0,1 0,3 88%
NICAM 3, Rec. J.17 du CCITT	14 bits linéaires Rec. J.17 du CCITT	-0,1 0,7 18%	-0,1 0,8 19%	-22 0,8 0,001%	0,1 1,0 60%

\bar{X} : opinion moyenne sur une échelle - 3 à + 3 conforme à la Recommandation 562.

S : écart type.

% Conf : niveau de confiance statistique indiquant que l'on préfère le codec «A» au codec «B».

Comme on peut le constater sur le tableau, il n'y a pas de différence significative entre les deux méthodes de codage lorsque l'ADM est échantillonnée à la fréquence de 330 kHz, sauf dans le cas du gong basse fréquence où une mise en forme du spectre de bruit est perçue. Cela a été principalement constaté au cours du temps d'amortissement du gong. Lorsque l'ADM est échantillonnée à la fréquence de 204 kHz, on peut également percevoir une dégradation dans la reproduction du gong à haute fréquence, donnant une qualité moins bonne que le NICAM 3. Pour ces deux cas, toutefois, on ne constate pas de différence significative entre les deux codecs pour des contenus de programme normaux.

2.2 Evaluation de qualité d'un codeur différentiel quasi instantané et d'un codeur MIC quasi instantané

On a procédé au Canada à des mesures objectives et à des essais subjectifs de deux schémas de codage proposés dans le Rapport 1075 (MOD I) pour la radiodiffusion de TVHD par satellite, à savoir:

- MICD quasi instantanée avec compression-extension (réduction de 15 à 8), proposé pour le système MUSE;*
- Compression-extension quasi instantanée (réduction de 14 à 10) comme spécifiée dans la Recommandation 651.

* Le schéma MICD quasi instantanée utilisé pour MUSE repose sur un coefficient de fuite de 0,9375, au lieu du coefficient de 0,975 qui a servi aux simulations évoquées ici. Les simulations avec des coefficients de fuite compris entre 0,95 et 0,99 ont fait apparaître seulement de légères différences du rapport signal/bruit.

2.2.1 Evaluation objective

Les deux schémas de codage ont été simulés sur ordinateur selon les descriptions données dans le Rapport 953-1 pour la compression-extension quasi instantanée et dans le Rapport 1075 pour la MICD quasi instantanée. Un coefficient de fuite de 0,975 a été utilisé dans la MICD quasi instantanée simulée. Les signaux son d'essai suivants ont été appliqués à chaque codeur:

- a) signaux sinusoïdaux à fréquence variable;
- b) sons synthétiques basse et haute fréquence;
- c) harpe naturelle.

Les signaux d'essai introduits dans les codeurs se composaient d'échantillons audio numériques de 32 kHz uniformément quantifiés à 15 bits/échantillon pour le codeur MICD quasi instantanée et à 14 bits/échantillon pour le codeur à compression-extension quasi instantanée. Pour chaque signal d'essai et pour chaque codeur, on a calculé et tracé, en fonction du temps des rapports signal/bruit sur des blocs successifs de 256 échantillons (8 ms).

Le signal sinusoïdal à fréquence variable se composait d'une séquence de dix sinusoïdes d'amplitude constante et de fréquence variable comprise entre 30 et 15 360 Hz. La Figure 4 a) donne un exemple de ce signal d'essai et les courbes du rapport signal/bruit obtenues sont tracées sur les Figures 4 b) et 4 c). Cette dernière figure montre qu'un schéma de quantification instantanée comme la compression-extension quasi instantanée se traduit par un rapport signal/bruit indépendant de la fréquence. Un schéma de codage par transition comme la MICD quasi instantanée donne de meilleures valeurs de rapport signal/bruit aux faibles fréquences (jusqu'à 1 kHz). En revanche, les valeurs de ce rapport se dégradent aux fréquences plus élevées, jusqu'à être inférieures à celles obtenues avec la compression-extension quasi instantanée.

Le signal son synthétique d'essai se composait des quatre premières notes de "Frère Jacques" (DO-RE-MI-DO) synthétisées à basse fréquence (60 - 80 Hz) et à plus haute fréquence (4 - 6 kHz). La Figure 5a) montre la forme du signal qui en résulte dans le domaine temporel; les courbes du rapport signal/bruit obtenu pour le son basse fréquence sont tracées sur la Figure 5b). La supériorité du codeur différentiel (MICD quasi instantanée) apparaît pour un signal basse fréquence de niveau élevé, malgré le moindre nombre de bits par échantillon (8 contre 10 pour la compression-extension quasi instantanée). Les deux codeurs donnent lieu à une décroissance du rapport signal/bruit à mesure que le niveau du signal diminue. Les courbes du rapport signal/bruit obtenues avec le son synthétique haute fréquence sont tracées sur la Figure 5c). Les valeurs du rapport signal/bruit obtenues pour la compression-extension quasi instantanée sont supérieures, pour des signaux de niveau élevé, à celles obtenues avec la MICD quasi instantanée.

Le son de harpe naturelle, représenté à la Figure 6a) révèle une très large gamme dynamique ainsi que de fortes variations de niveau sur de courtes périodes. Les courbes du rapport signal/bruit obtenues avec ce signal d'essai font l'objet des Figures 6b) et 6c). Ces deux courbes sont comparables si ce n'est, peut-être, un léger avantage pour la MICD quasi instantanée par rapport à la compression-extension quasi instantanée lorsque le signal d'entrée atteint la valeur crête de sa gamme dynamique.

2.2.2 Evaluation subjective

Une série d'essais officieux pour lesquels on a utilisé des casques d'écoute a été effectuée avec 4 ingénieurs du son et 2 élèves-ingénieurs comme auditeurs. Les essais ont été faits selon la méthode de comparaison proposée par le CCIR (Recommandation 562-2), dans laquelle la séquence A-B est présentée deux fois, après quoi l'auditeur dispose d'environ 15 s pour donner une note de -3 (codeur A nettement plus mauvais que le codeur B) à +3 (codeur A bien meilleur que codeur B).



Pour les essais on a pris des extraits d'un solo de harpe, un solo d'orgue, un solo féminin et des morceaux joués par un ensemble de musique de chambre, par un quatuor de jazz et un orchestre de musique pop. Les sons synthétiques basse et haute fréquence décrits au § 2.2.1 ont été aussi inclus dans les essais. La loi de préaccentuation/désaccentuation de 50/15 μ s décrite dans la Recommandation 651 a été appliquée aux deux codeurs MICD quasi instantanée et compression-extension quasi instantanée pour les essais d'écoute.

Bien qu'officieux, les essais d'écoute ont donné des résultats cohérents. Les auditeurs n'ont pas constaté une grande différence entre les codeurs MICD quasi instantanée et compression-extension quasi instantanée pour les six extraits de musique naturelle (à savoir les solos de harpe, d'orgue et de voix de femme, et les extraits de musique de chambre, de jazz et de musique pop). La note moyenne d'opinion pour ces six fréquences a été de 0,2 en faveur de la MICD quasi instantanée avec un écart type de 0,8.

Le groupe d'auditeurs a nettement préféré le codeur MICD quasi instantanée pour les sons synthétiques à basse fréquence (note moyenne 2,2, écart type 0,7) et la compression-extension quasi instantanée pour les sons synthétiques de fréquence plus élevée (note, moyenne de -1,8 avec un écart type de 0,7). Ces résultats concordent avec les résultats des mesures objectives du § 2.2.1, où le codeur MICD quasi instantanée a permis d'obtenir, par rapport à la compression-extension quasi instantanée, de meilleures valeurs du rapport signal/bruit pour les signaux de basse fréquence mais de moins bonnes valeurs de ce rapport aux fréquences plus élevées.

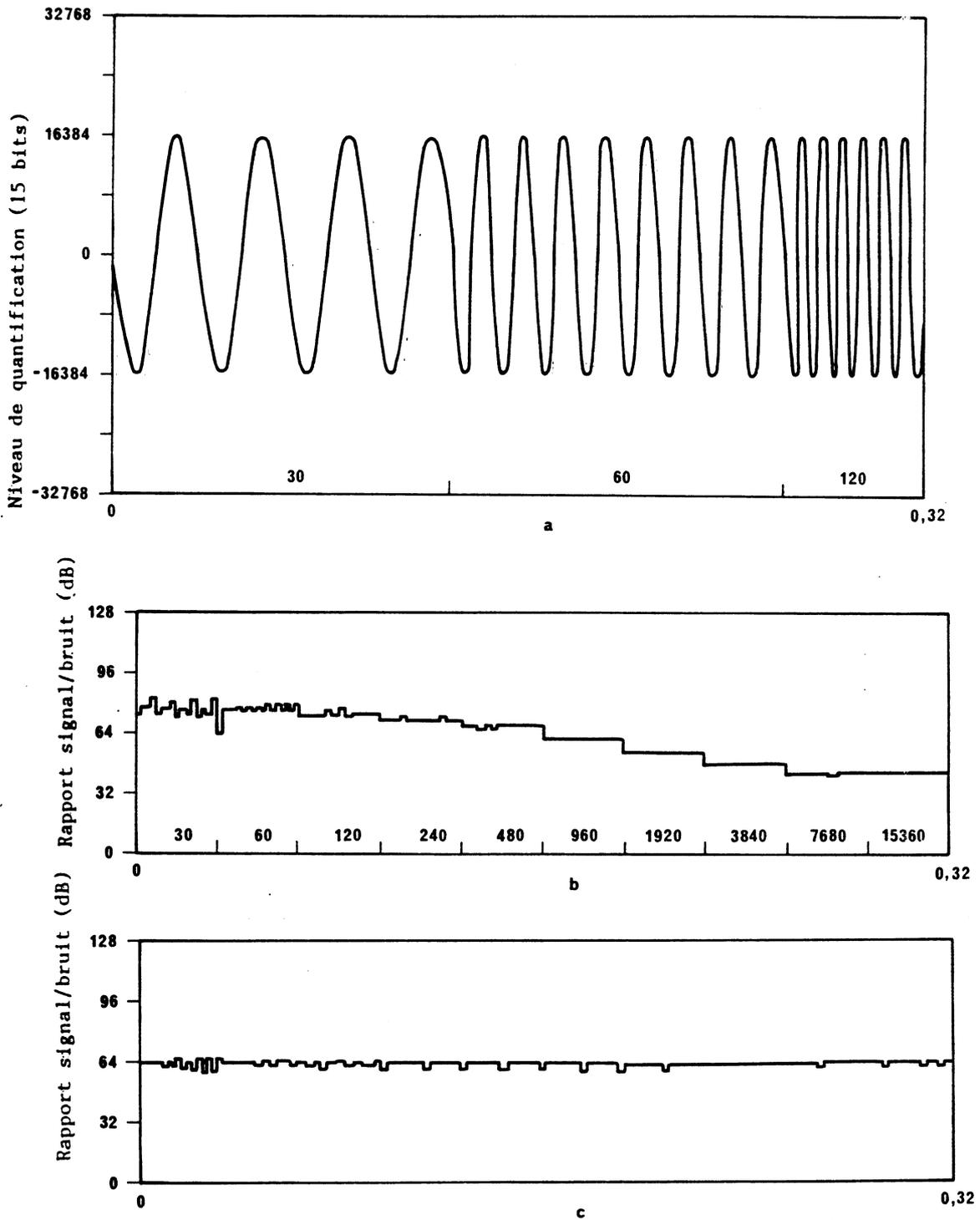


FIGURE 4

Signal sinusoidal à fréquence variable

- a) exemple de signal d'essai
- b) MICD quasi instantanée: rapport signal/bruit en fonction de la fréquence de l'onde sinusoidale
- c) compression-extension quasi instantanée: rapport signal/bruit en fonction de la fréquence de l'onde sinusoidale

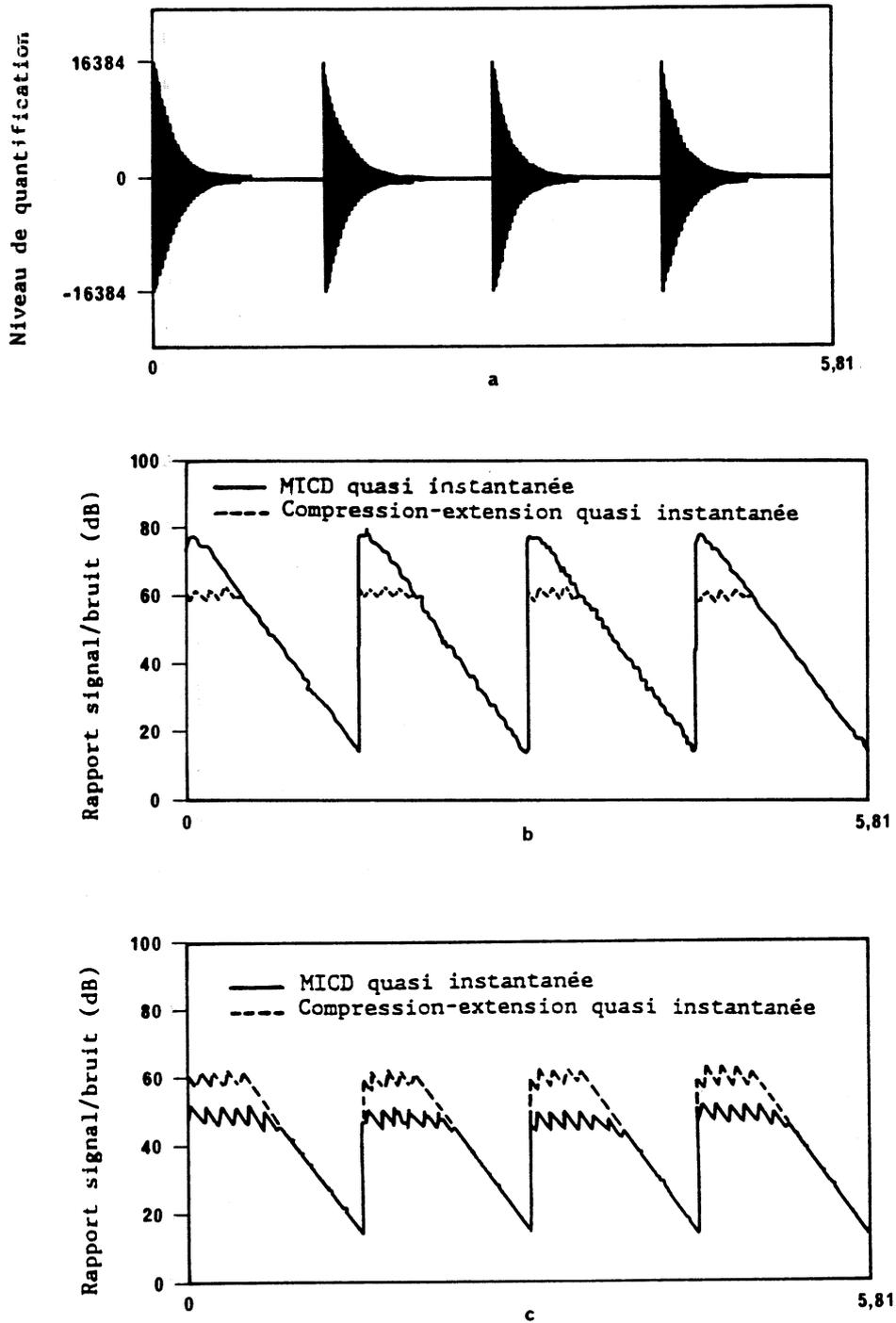


FIGURE 5

Sons synthétiques de basse et de haute fréquence

- a) signal d'essai
- b) courbes du rapport signal/bruit pour le son basse fréquence
- c) courbes du rapport signal/bruit pour le son haute fréquence

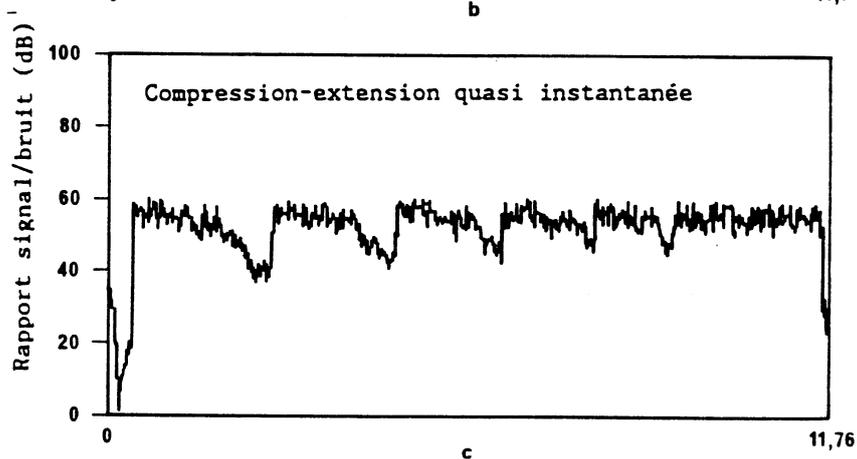
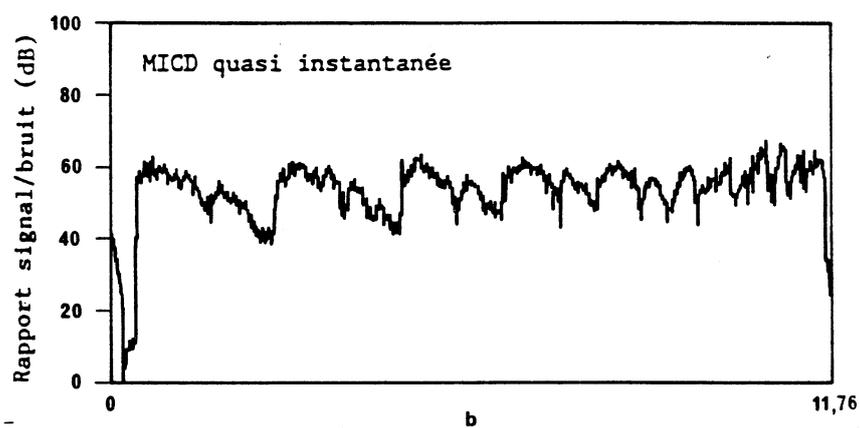
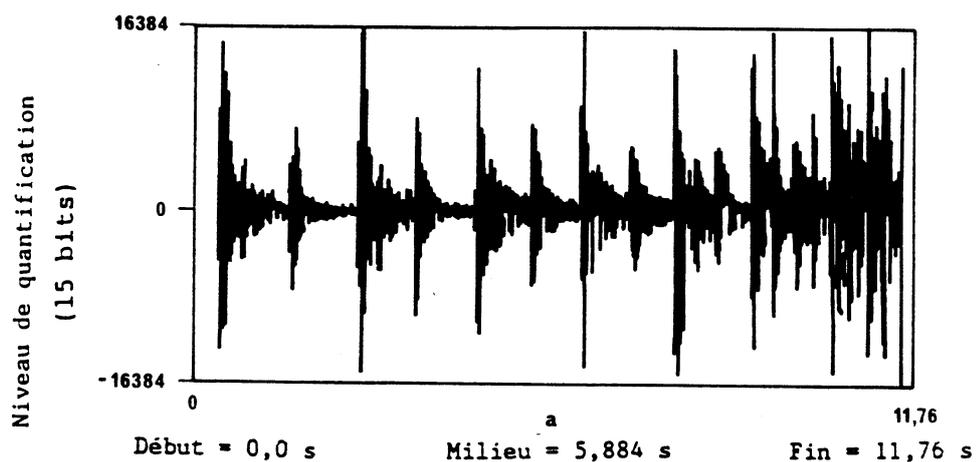


FIGURE 6

Son naturel de harpe

- a) signal d'essai
- b) courbe du rapport signal/bruit pour le codeur MICD quasi instantanée
- c) courbe du rapport signal/bruit pour le codeur compression-extension quasi instantanée

3. Mesures subjectives effectuées aux Etats-Unis

On a effectué, avec des auditeurs entraînés, des essais d'écoute soigneusement contrôlés, pour évaluer le système à modulation delta adaptative que décrit le Rapport 953-1. Ces essais consistaient à tester le système à MDA (fréquence d'échantillonnage = 300 kHz) et à le comparer au disque compact à grande dynamique.

On a engagé pour ces essais 24 auditeurs très motivés et hautement qualifiés dont l'ouïe, selon une mesure audiométrique effectuée par un audiologiste diplômé, était cliniquement normale. Ils étaient âgés de 16 à 27 ans, ce qui garantissait une bonne probabilité d'ouïe normale. Tous ces auditeurs expérimentés et experts étaient invités à "voter" pour le système dont la qualité sonore leur paraissait la meilleure. On ne leur précisait pas ce qu'ils allaient écouter. Le groupe se composait de 9 femmes et 15 hommes, et on les rémunérait pour augmenter leur motivation.

Les conditions d'écoute étaient très proches de l'environnement que stipule la Recommandation 562 de la Commission d'études 10. Le studio où avaient lieu les essais était revêtu de moquette, tendu de rideaux, avait un temps de réverbération d'environ 0,25 seconde aux médiums et mesurait 4,9 x 6,7 m. Les haut-parleurs étaient séparés par une distance de l'ordre de 3 m. Les auditeurs participaient aux essais un par un.

Pour les passages musicaux les plus bruyants, on fixait à 85 dB le niveau sonore au centre du studio. Ce niveau était retenu car à des niveaux plus faibles certaines différences entre les systèmes n'auraient peut-être pas été perceptibles.

Les 10 extraits choisis sur disques compacts (voir le Tableau III) étaient tirés d'un disque compact (DC) avec matériel d'essai de la qualité des signaux audio (SQAM) de l'Union européenne de radiodiffusion (UER) et servaient d'extraits critiques pour les essais:

TABLEAU III

Extraits de programmes

Musique électronique - Frère Jacques ABBA - Musique pop Triangles - Instrument isolé Piano de concert - Notes séparées et courte sélection musicale Cloches - Instrument isolé Eddie Rabbitt - "Country music" Soprano - Solo Ensemble de vents - Courte sélection musicale Xylophone - Notes séparées et sélection musicale Parole, homme - En allemand

Ces sélections stéréophoniques qui servaient aux essais présentent une large gamme de dynamiques, de spectre et de caractéristiques temporelles. Elles représentent des extraits de programmes, typiques ou non. On a choisi les plus éprouvants pour le type de traitement à MDA essayé et pour révéler les défauts introduits par le codage. Les défauts qui pourraient être audibles sont: la linéarité A/N et N/A, la distorsion par repliement du spectre et le bruit modulé par le programme.

Les extraits duraient de 20 à 140 secondes et l'essai subjectif en tout une demi-heure.

On utilisait la méthode psychophysique classique du choix forcé à tiers exclu (two-alternative forced-choice, 2 AFC). (La méthode de la Recommandation 562 de la Commission d'études 10 ne convenait pas car elle repose sur une détection du signal, non sur une notation de la qualité.) Les dix sélections étaient présentées quatre fois chacune, dans deux séquences aléatoires. Les auditeurs passaient alternativement de l'un à l'autre système puis, dès la fin de chaque extrait, votaient pour celui dont la qualité sonore était supérieure. Cinq secondes environ séparaient les extraits.

Le Tableau [VI] présente les résultats des essais initiaux; ils indiquent pour chaque extrait le nombre d'appréciations et le pourcentage correspondant. On peut ainsi voir si le choix de l'extrait a une influence sur le jugement de l'auditeur et comparer les deux voies son. Lorsque la différence est imperceptible, on obtient des résultats distribués de façon aléatoire avec une moyenne $N/2$ (50%) et un écart type en racine carrée de $N/4$ (N étant le nombre d'essais).

Chaque auditeur formulait 40 appréciations, soit 4 votes pour chacun des dix extraits. Ainsi, pour tous les extraits les 24 auditeurs formulaient 960 appréciations. Une sélection aléatoire donnerait 48 votes (50%) par extrait en faveur du disque compact, avec un écart type de 4,9 votes (5,1%).

Le résultat global des essais initiaux (Tableau IV) donne une préférence de 52,5% pour la MDA et 47,5% pour le disque compact, ce qui représente une très faible différence, si même il y en a une. On peut conclure de cet examen global que le compresseur-extenseur est transparent du point de vue acoustique. Il y a toutefois 5 extraits pour lesquels on a marqué une préférence différant de plus de 5% de la moyenne 50%: la musique électronique, 55,2% en faveur du disque compact; ABBA, 65,6% pour la MDA; les triangles, 61,5% pour le disque; le solo de soprano, 59,4% pour la MDA; et le xylophone, 62,5% pour la MDA; il semble que 3 à 5 des auditeurs aient été les principaux responsables de ces écarts.

Comme quelques auditeurs étaient capables de détecter une différence pour ces extraits critiques (bien qu'ils n'aient exprimé aucune préférence pour l'une ou l'autre des voies son), on a procédé à une étude de suivi pour voir si les résultats n'étaient pas dus à de légères différences de niveau audio. L'adaptation de niveau des deux voies (disque en direct et disque via le codex MDA) a été améliorée d'environ 1 dB à mieux que 0,5 dB. Les autres conditions des essais restaient inchangées. Le Tableau V donne les résultats.

A la suite de ces nouvelles adaptations du niveau, les auditeurs experts ont eu encore plus de mal à distinguer le codec MDA du disque en direct. Pour trois des cinq extraits de programme, les résultats se sont déplacés d'environ 10 points en pourcentage vers une répartition purement aléatoire (50/50) et on a abouti aux pourcentages 44, 50 et 52%. Pour les deux autres extraits, les résultats n'ont pas varié de façon appréciable (2 à 3 points en pourcentage).

Un essai soigneusement contrôlé dans les conditions ci-dessus a montré que des auditeurs experts et expérimentés n'exprimaient aucune préférence pour le disque compact soit en direct, soit transmis via un codec MDA."

TABLEAU IV

Principaux résultats des essais d'écoute
Août 1988

Votes en faveur du lecteur de disque compact
24 auditeurs
96 jugements à choix forcé par extrait

	Nombre total de votes	Pourcentage
1. Musique électronique # 7	53	55,2
2. ABBA # 69	33	34,4
3. Triangles # 32	59	61,5
4. Piano # 39	47	49,0
5. Cloches # 34	48	50,0
6. Eddie Rabbitt # 70	49	51,0
7. Soprano # 44	39	40,6
8. Ensemble à vent # 66	47	49,0
9. Xylophone # 36	36	37,5
10. Parole, homme (en allemand) # 54	45	47,0
Totaux	456	47,5

Pour 96 événements aléatoires, l'écart type = 4,9 = 5,1%

TABLEAU V

Résultat des essais d'écoute de suivi
Septembre 1988

Votes en faveur du lecteur de disque compact
20 auditeurs
160 jugements à choix forcé par extrait

	Nombre total de votes	Pourcentage
1. Musique électronique # 7	86	53,8
2. ABBA # 69	70	43,8
3. Triangles # 32	80	50,0
4. Soprano # 44	83	52,0
5. Xylophone # 36	65	40,6
Totaux	384	48,0

Pour 160 événements aléatoires, l'écart type = 6,3 = 4%

RÉFÉRENCES BIBLIOGRAPHIQUES

AuBC [mai 1985] Australian Broadcasting Corporation Engineering Research and Development Report No. 139. Subjective Tests of Adaptive Delta Modulation With B-MAC for HACBSS.