

A/85:2011
26 juin 2011

Pratique recommandée par l'ATSC : Techniques d'établissement et de maintien de l'intensité sonore pour la télévision numérique

Advanced Television Systems Committee
1776 K Street, N.W.
Suite 200
Washington, D.C. 20006
www.atsc.org

L'ATSC (Advanced Television Systems Committee, Inc.) est un organisme international sans but lucratif, qui élabore des normes volontaires visant la télévision numérique. Les organismes membres de l'ATSC représentent les industries de la radiodiffusion, du matériel de radiodiffusion, du cinéma, de l'électronique grand public, de l'informatique, de la câblodistribution, des satellites et des semi-conducteurs.

Plus particulièrement, l'ATSC s'efforce de coordonner les normes de télévision entre les divers médias de communication en mettant l'accent sur la télévision numérique, sur les systèmes interactifs et sur les communications multimédias large bande. L'ATSC élabore aussi des stratégies de mise en œuvre de la télévision numérique et présente des séminaires de sensibilisation sur ses normes.

L'ATSC a été formé en 1982 par les organismes membres du Joint Committee on InterSociety Coordination (JCIC) : l'Electronic Industries Association (EIA), l'Institute of Electrical and Electronic Engineers (IEEE), le National Association of Broadcasters (NAB), la National Cable Telecommunications Association (NCTA) et la Society of Motion Picture and Television Engineers (SMPTE). Actuellement, il compte environ 150 membres, qui représentent les industries de la radiodiffusion, du matériel de radiodiffusion, du cinéma, de l'électronique grand public, de l'informatique, de la câblodistribution, des satellites et des semi-conducteurs.

Les normes ATSC sur la télévision numérique portent sur la télévision haute définition (TVHD), sur la télévision de définition standard (TVDS), sur la radiodiffusion de données, sur l'audio ambiophonique multicanal et sur la radiodiffusion directe par satellite.

Note : Nous attirons l'attention du lecteur sur la possibilité que la conformité à la présente pratique recommandée nécessite l'utilisation d'une invention faisant l'objet de droits de brevet. La publication de ce document ne constitue pas une prise de position à l'égard de la validité de cette allégation ou des droits de brevets pouvant y être liés. Toutefois, un ou plusieurs titulaires de brevets ont pu déposer une déclaration concernant les modalités suivant lesquelles le ou les titulaires de brevets consentiraient à accorder une licence d'utilisation en vertu de ces droits à des personnes ou à des entités souhaitant acquérir une telle licence. Les détails peuvent être obtenus auprès du secrétariat de l'ATSC et du titulaire de brevet.

Table des matières

1	PORTÉE.....	9
1.1	Contexte et introduction	9
1.2	Structure	10
2	RÉFÉRENCES.....	11
2.1	Références informatives	11
3	DÉFINITION DES TERMES.....	13
3.1	Indication de conformité	13
3.2	Traitement des éléments syntaxiques	14
3.3	Sigles et abréviations	14
3.4	Termes	14
4	LE SYSTÈME AUDIO MULTICANAL AC-3.....	16
5	MESURE DE L'INTENSITÉ SONORE	16
5.1	Aperçu général de la méthode de mesure de l'intensité sonore UIT-R BS.1770	17
5.2	Mesure	18
5.2.1	Mesure durant la production ou la postproduction	18
5.2.2	Mesure en temps réel en cours de production (événement en direct)	18
5.2.3	Mesure de contenu de longue durée fini	19
5.2.4	Mesure de contenu de courte durée	19
5.2.5	Mesure fondée sur des fichiers	19
6	NIVEAU SONORE CIBLE ET NIVEAU DE CRÊTE VRAI POUR LA DISTRIBUTION OU L'ÉCHANGE DE CONTENU.....	19
7	POINTS À CONSIDÉRER DANS LA GESTION DES MÉTADONNÉES AYANT UN EFFET SUR L'INTENSITÉ SONORE DES PROGRAMMES AUDIO	20
7.1	Importance de dialnorm	20
7.2	Modes de gestion des métadonnées	21
7.3	Utilisation de métadonnées dialnorm statiques	21
7.3.1	Établissement de la valeur dialnorm par la méthode de la moyenne longue durée	22
7.3.2	Établissement de la valeur dialnorm pour la production	22
7.3.3	Contenu non conforme à l'Intensité sonore cible	22
7.3.4	Valeur dialnorm et contrôle de qualité d'intensité sonore	22
7.3.5	Réglage dialnorm d'émission pour la conformité à la norme A/53	22
7.3.6	Avantage du paramètre dialnorm statique	23
7.4	Utilisation de métadonnées dialnorm prédéfinies	23
7.4.1	Mise en œuvre	24
7.4.2	Exigence de synchro trame AC-3	24
7.5	Utilisation de métadonnées dialnorm dynamiques	24
7.5.1	Déploiement du système	25
7.5.1.1	Métadonnées Dolby E sur liaison série	26
7.5.1.2	Métadonnées Dolby E en VANC	26
7.5.1.3	Métadonnées et codecs	26
7.5.1.4	Métadonnées enregistrées dans des fichiers	26
7.5.2	Technique de production en direct	26
7.5.3	Technique de production en différé	26
7.5.4	Contrôle de la production	27

7.5.5	Métadonnées semi-dynamiques	27
7.5.6	Incidence de la perte de métadonnées sur le contenu	27
7.5.7	Hybride statique-dynamique	27
7.5.8	Avantages des métadonnées dynamiques	28
8	MÉTHODES DE CONTRÔLE EFFICACE DE L'INTENSITÉ SONORE ENTRE PROGRAMME ET CONTENU INTERSTITIEL	28
8.1	Solutions efficaces	28
8.1.1	Pour les opérateurs employant un système à dialnorm statique (voir la section 7.2)	28
8.1.2	Pour les opérateurs employant un système dialnorm dynamique (voir la section 7.5)	28
8.2	Conditions hostiles	29
8.3	Recommandations sommaires	29
8.4	Insertion de publicité locale par la station de télévision et le MVPD	30
9	GESTION DE LA DYNAMIQUE	30
9.1	Système de contrôle de dynamique AC-3 (« réversible »)	31
9.1.1	DRC en mode ligne et en mode RF	31
9.1.2	Contrôle	32
9.1.3	Relations avec dialnorm	32
9.1.4	Codage professionnel	33
9.1.5	DRC AC-3 : sélection de « None »	33
9.2	Traitement de la dynamique avec interface de métadonnées (« réversible »)	34
9.3	Traitement de la dynamique sans interface de métadonnées (irréversible)	35
9.4	Expérience du consommateur	36
10	CONFIGURATION DE CONTRÔLE AUDIO	36
10.1	Contexte	37
10.2	Caractéristiques des locaux et des espaces	37
10.3	Installation	38
10.4	Étalonnage du niveau de référence	39
10.5	Détails des signaux d'essai	42
10.5.1	Définition de 0 dB FS	42
10.5.2	Tonalité sinusoïdale 440 Hz	42
10.5.3	Bruit rose limité à la bande	42
	ANNEXE A: INTENSITÉ SONORE DE PROGRAMME	44
A.1	INTRODUCTION	44
A.2	INTENSITÉ SONORE DES PROGRAMMES	44
A.3	CRÊTE VRAIE	48
	ANNEXE B: ACOUSTIQUE DE SALLE ET DISPOSITION DES HAUT-PARLEURS.....	50
	ANNEXE C: CORRECTION ACOUSTIQUE.....	52
	ANNEXE D: AIDE-MÉMOIRE DE CONFIGURATION DE CONTRÔLE DE RÉFÉRENCE POUR LA TÉLÉVISION	
	53	
	ANNEXE E: PLAGE D'INTENSITÉ SONORE.....	55
	ANNEXE F : DÉTAILS SUR LE CONTRÔLE DE GAMME DYNAMIQUE AC-3	57
F.1	APERÇU GÉNÉRAL DU DRC	57
F.2	PLACEMENT DES MÉTADONNÉES DANS LES FLUX BINAIRES DRC AC-3	59

F.3	CALCULS DES MOTS DE GAIN AC-3	60
F.4	PONDÉRATION D'INTENSITÉ SONORE ET COURBES D'ENTRÉES-SORTIE DE DRC	65
F.5	PARAMÈTRES D'ENCODEUR POUR LE RÉGLAGE DES MÉTADONNÉES DRC	66
F.6	COMMANDES EXÉCUTABLES AU DÉCODEUR DRC	67
ANNEXE G: PARAMÈTRES DE DONNÉES AC-3		68
ANNEXE H: AIDE-MÉMOIRE À L'INTENTION DES TECHNICIENS DE STATION ET DE MVPD GESTION DE L'INTENSITÉ SONORE		70
H.1	INTRODUCTION	70
H.2	PORTÉE	70
H.3	DÉFINITIONS	70
H.4	GESTION DE L'INTENSITÉ SONORE	71
H.5	EXIGENCE DE LA FCC	71
H.6	MESURE DU CONTENU DISTRIBUÉ	71
	H.6.1 Contenu longue durée	71
	H.6.2 Contenu courte durée	72
	H.6.3 Bulletins de nouvelles ou autre programmation en direct	72
	H.6.4 Contenu enregistré dans des fichiers	72
H.7	INTENSITÉ SONORE CIBLE POUR FACILITER L'ÉCHANGE DE PROGRAMMES	72
H.8	MÉTHODES DE CONTRÔLE EFFICACE DE L'INTENSITÉ SONORE ENTRE PROGRAMME ET CONTENU INTERSTITIEL	73
H.9	RÉGLAGE dialnorm PAR UNE STATION AFFILIÉE	73
H.10	INSERTION DE CONTENU PAR LA STATION DE TÉLÉVISION OU LE MVPD	74
H.11	CONTRÔLE DE GAMME DYNAMIQUE (DRC) AC-3	74
ANNEXE I: AIDE-MÉMOIRE À L'INTENTION DES MIXEURS ET DES MONTEURS SON QUI CRÉENT DU CONTENU		75
I.1	INTRODUCTION	75
I.2	PORTÉE	75
I.3	DÉFINITIONS	75
I.4	ENVIRONNEMENT DE CONTRÔLE	76
I.5	CONTRÔLE DE NIVEAU UIT-R BS.1770	77
I.6	INTENSITÉ SONORE DE CONTENU DURANT LE MIXAGE	77
I.7	INTENSITÉ SONORE CIBLE POUR LE CONTENU SANS MÉTADONNÉES	77
I.8	EXIGENCE DE LA FCC	77
I.9	MESURE DE CONTENU POSTPRODUIT	78
	I.9.1 Contenu longue durée	78
	I.9.2 Contenu courte durée	78
ANNEXE J: EXIGENCES RELATIVES À L'ÉTABLISSEMENT ET AU MAINTIEN DE L'INTENSITÉ SONORE DES MESSAGES PUBLICITAIRES À LA TÉLÉVISION NUMÉRIQUE		79
J.1	INTRODUCTION ET PORTÉE	79
J.2	GESTION DE L'INTENSITÉ SONORE	79
J.3	PARAMÈTRE dialnorm POUR LES SYSTÈMES AC-3	79
J.4	MESURE DE L'INTENSITÉ SONORE DES MESSAGES PUBLICITAIRES	79
J.5	MESSAGE PUBLICITAIRES AU POINT D'INSERTION	79

ANNEXE K: EXIGENCES RELATIVES À L'ÉTABLISSEMENT ET AU MAINTIEN DE L'INTENSITÉ SONORE DES MESSAGES PUBLICITAIRES À LA TÉLÉVISION NUMÉRIQUE QUAND ON UTILISE DES CODECS NON-AC-3 80

K.1.	INTRODUCTION ET PORTÉE	80
K.2.	GESTION DE L'INTENSITÉ SONORE	80
K.3.	INTENSITÉ SONORE DES CANAUX NON-AC-3	80
K.4.	MESURE DE L'INTENSITÉ SONORE DES MESSAGES PUBLICITAIRES	80
K.5.	MESSAGE PUBLICITAIRE AU POINT D'INSERTION	80

Index des tableaux et des figures

Tableau	10.1	Catégories de régies du son utilisées en production télévisuelle	Err
		or! Bookmark not defined.	
Tableau	10.2	Niveau de pression acoustique de référence	Err
		or! Bookmark not defined.	
Tableau D.1		Niveau de pression acoustique de référence	51
Tableau G.1		Métadonnée de contrôle critique	66
Tableau G.2		Métadonnée de contrôle de base - Encodeur	66
Tableau G.3		Métadonnée de contrôle de base - Décodeur	66
Tableau G.4		Métadonnée informationnelle	67
Tableau 10.2		Niveau de pression acoustique de référence	74
Figure 5.1		Algorithme de mesure de l'intensité sonore de l'UIT-R BS.1770	17
Figure 7.1		Concept de métadonnées statiques	21
Figure 7.2		Concept de métadonnées prédéfinies	23
Figure 7.3		Concept de métadonnées dynamiques	25
Figure 7.4		Options d'insertion des métadonnées	27
Figure 10.1		Courbe de réponse électroacoustique dans une salle fonctionnelle	Err
		or! Bookmark not defined.	
Figure A.1		Schéma fonctionnel de l'algorithme de mesure de l'intensité sonore multicanal	Err
		or! Bookmark not defined.	
Figure	A.2	Courbe de pondération	RLB Err
		or! Bookmark not defined.	
Figure A.3		Réponse du préfiltre utilisé pour tenir compte des effets acoustiques de la tête	Err
		or! Bookmark not defined.	
Figure A.4		Résultats regroupés pour les trois ensembles de données ($r = 0,977$).	Err
		or! Bookmark not defined.	
Figure A.5		Niveau de crête du signal continu et du signal échantillonné	Err
		or! Bookmark not defined.	
Figure A.6		Structure de base de l'algorithme d'estimation de la crête vraie défini dans UIT-R BS.1770-1	

								Err
		or! Bookmark not defined.						
Figure	E.1	Seuils	d'intensité	sonore	critiques			Err
		or! Bookmark not defined.						
Figure	F.1	Principe de base du contrôle de gamme dynamique (DRC)						AC-3 Err
		or! Bookmark not defined.						
Figure	F.2	Exigences relatives au décodeur						AC-3 (mode ligne) Err
		or! Bookmark not defined.						
Figure	F.3	Exigences relatives au décodeur						AC-3 (mode RF). Err
		or! Bookmark not defined.						
Figure	F.4	Placement des métadonnées de DRC						AC-3 dans le flux binaire Err
		or! Bookmark not defined.						
Figure	F.5	Calculs des mots de gain						AC-3 par l'encodeur (partie 1). Err
		or! Bookmark not defined.						
Figure	F.6	Calculs des mots de gain						AC-3 par l'encodeur (partie 2) Err
		or! Bookmark not defined.						
Figure	F.7	Lissage du gain par la méthode "chevauchement & addition"						Err
		or! Bookmark not defined.						
Figure	F.8	Courbe de pondération						d'intensité sonore DRC Err
		or! Bookmark not defined.						
Figure	F.9	Caractéristiques d'entrée-sortie						des profils DRC Err
		or! Bookmark not defined.						

Pratique recommandée par l'ATSC :

Techniques d'établissement et de maintien de l'intensité sonore pour la télévision numérique

1 PORTÉE

La présente pratique recommandée (PR) de l'ATSC fournit des conseils aux radiodiffuseurs et aux créateurs de bandes sonores pour le contenu télévisuel ATSC haute définition (HD) et de définition standard (DS). Elle recommande les méthodes de production, de distribution et de transmission nécessaires pour offrir les pistes sonores de la plus haute qualité aux téléspectateurs du numérique.

Cette PR est centrée sur les techniques de mesure, de production et de contrôle postproduction du son et des méthodes permettant de contrôler avec efficacité l'intensité sonore dans la distribution ou l'échange de contenu. Elle recommande les méthodes visant à contrôler efficacement l'intensité sonore entre programme et contenu interstitiel, explique les systèmes de métadonnées et leur utilisation et décrit le contrôle moderne de la gamme dynamique. Cette PR contient aussi des renseignements particuliers sur la gestion de l'intensité sonore aux frontières entre les programmes et le contenu interstitiel.

En outre, la PR est accompagnée de deux annexes pouvant être détachées et servir d'aide-mémoire à deux communautés d'intérêts en particulier : ingénieurs de station/distributeur multicanal de programmes audiovisuels (MVPD) et personnel de production sonore.

1.1 Contexte et introduction

La pratique recommandée A/85:2011, 25 juillet 2011 succède à la pratique recommandée A/85:2009 datant du 25 mai 2011.

Malgré la fin de la transition à la TVN, de nombreux radiodiffuseurs et la communauté de la production ont été lents à s'adapter aux changements nécessaires à la transition des techniques sonores NTSC analogiques aux pratiques audio numériques modernes. La gamme dynamique audible étendue (plus de 100 dB) de la télévision numérique ouvre la voie à une variation excessive du contenu sonore lorsque l'intensité sonore de la TVN n'est pas gérée de façon appropriée.

Les consommateurs ne s'attendent pas à un grand changement d'intensité sonore entre un programme et un contenu interstitiel ou quand ils passent d'une chaîne à une autre. L'utilisation abusive de la gamme dynamique étendue disponible a entraîné des plaintes des consommateurs, qui doivent garder leur télécommande à portée de la main et régler le volume pour le confort de l'écoute.

Le système de télévision analogique NTSC a recours à un traitement conventionnel de la gamme dynamique sonore à diverses étapes du trajet du signal afin de gérer l'intensité sonore pour les radiodiffusions. Cette méthode compense les limites de la gamme dynamique du matériel analogique et contrôle les divers niveaux d'intensité du signal sonore reçu des fournisseurs. Elle aide aussi à adoucir les passages des programmes aux contenus interstitiels. Malgré sa simplicité et son efficacité, cette méthode réduit en permanence la gamme dynamique et modifie les signaux sonores avant qu'ils atteignent le téléspectateur. Elle change les caractéristiques du son original, le modifiant par rapport à ce que le fournisseur de programmes avait prévu, pour l'adapter aux limites du système analogique.

Le système audio AC-3 défini dans la norme de télévision numérique ATSC fait appel aux métadonnées, c.-à-d. aux « données à propos des données », pour contrôler l'intensité sonore et les autres paramètres audio de façon plus efficace, sans modifier irrémédiablement la gamme dynamique du contenu. Le fournisseur de contenu ou le télédiffuseur numérique code les métadonnées dans le contenu sonore. Du point de vue des téléspectateurs, le paramètre de métadonnée *Dialog Normalization* (dialnorm) règle les différents contenus à une intensité sonore uniforme de manière transparente. Il permet d'atteindre un résultat semblable à l'utilisation de la télécommande par le téléspectateur qui règle le volume sonore à un niveau confortable en écoutant des programmes télé disparates ou au passage d'un programme à un message publicitaire ou d'une chaîne à une autre. Le paramètre dialnorm et les autres paramètres de métadonnées font partie intégrante du flux binaire audio AC-3.

Le document ATSC A/53, partie 5:2010 [1] rend obligatoire le transport du paramètre dialnorm et de valeurs dialnorm réglées de façon appropriée.

L'industrie reconnaît qu'une nouvelle compétence en matière de mesure de l'intensité sonore, de contrôle de la production, d'utilisation des métadonnées et de pratiques modernes visant la gamme dynamique est vitale pour répondre aux attentes du fournisseur de contenu, du radiodiffuseur, des téléspectateurs et des organismes de réglementation.

Le présent document contient des recommandations et des renseignements techniques relatifs aux éléments suivants :

- Mesure de l'intensité sonore suivant la recommandation UIT-R BS.1770.
- Intensité sonore cible pour l'échange de contenu sans métadonnées.
- Établissement d'environnements de contrôle de référence pour la production avec la gamme dynamique étendue inhérente à la télévision numérique, en prenant en compte les multiples environnements d'écoute à la maison.
- Présentation de méthodes efficaces pour contrôler l'intensité sonore lors de transitions entre programmes et contenus interstitiels.
- Utilisation efficace des métadonnées audio pour la production, la distribution et la transmission de contenu numérique.
- Contrôle de la gamme dynamique à l'intérieur des signaux audio AC-3 et contrôle moderne conventionnel de la gamme dynamique à titre d'ajout ou de solution de rechange, y compris des recommandations visant la gestion de l'intensité sonore et de la dynamique aux frontières entre les programmes et le contenu interstitiel.

1.2 Structure

La structure de ce document est la suivante :

- **Section 1** – Définition de la portée du document et introduction générale.
- **Section 2** – Liste des références et des documents pertinents.
- **Section 3** – Définition des termes, sigles et abréviations utilisés dans le document.
- **Section 4** – Explication du sommaire technique du système audio multicanal AC-3.
- **Section 5** – Explication de la mesure de l'intensité sonore fondée sur les techniques définies dans la recommandation UIT-R BS.1770.
- **Section 6** – Recommandations visant l'échange de contenu sans métadonnées.
- **Section 7** – Lignes directrices visant l'utilisation pratique de métadonnées dynamiques et statiques dans les environnements de production, de distribution et de transmission.

- **Section 8** – Description des méthodes de contrôle efficaces de l'intensité sonore lors de transitions entre programme et contenu interstitiel.
- **Section 9** – Examen des principaux problèmes liés au contrôle de la gamme dynamique (DRC).
- **Section 10** – Indication de la configuration des systèmes audio pour la télévision numérique, y compris l'alignement des haut-parleurs de contrôle de la régie à un niveau de pression acoustique de référence.
- **Annexe A** – Intensité sonore des programmes : données contextuelles sur les algorithmes de mesure de l'intensité sonore et des niveaux de crête vrais de la recommandation BS.1770.
- **Annexe B** – Acoustique des salles et disposition des haut-parleurs : explication des principes de base du contrôle (*monitoring*) dans la régie.
- **Annexe C** – Correction de la salle : examen des problèmes liés à l'interaction entre le son provenant des haut-parleurs et la salle.
- **Annexe D** – Aide-mémoire sur le réglage du niveau acoustique de référence pour les salles de contrôle de télévision.
- **Annexe E** – Plages d'intensité sonore : examen de la plage d'intensité sonore à l'intérieur de laquelle un auditeur accepte les variations d'intensité sonore à l'intérieur d'éléments de contenu et entre des éléments de contenu.
- **Annexe F** – Détails relatifs au contrôle de la gamme dynamique AC-3.
- **Annexe G** – Paramètres de métadonnées AC-3.
- **Annexe H** – Aide-mémoire à l'intention des ingénieurs de station et de MVPD pour la gestion de l'intensité sonore par les stations et les opérateurs de MVPD.
- **Annexe I** – Aide-mémoire à l'intention du personnel de mélange audio qui crée du contenu (messages publicitaires et programmes).
- **Annexe J** – Exigences relatives à l'établissement et au maintien de l'intensité sonore des messages publicitaires en télévision numérique.
- **Annexe K** – Exigences relatives à l'établissement et au maintien de l'intensité sonore des messages publicitaires en télévision numérique au moyen de codecs audio non-AC-3.

2 RÉFÉRENCES

Les éditions indiquées ci-dessous sont valides au moment de la publication du document. Les documents cités en référence faisant tous l'objet de révisions, les utilisateurs de la présente pratique recommandée sont invités à vérifier s'il est possible d'appliquer une édition plus récente du document cité.

2.1 Références informatives

Les documents suivants contiennent des renseignements pouvant être utiles dans l'application de la présente pratique recommandée.

- [1] ATSC. *ATSC Digital Television Standard, Part 5 – AC-3 Audio System Characteristics*, Doc. A/53 Part 5:2010, Advanced Television Systems Committee, Washington, D.C., 6 juillet 2010.
- [2] IEEE/ASTM SI 10-2002, *Use of the International Systems of Units (SI): The Modern Metric System*, Institute of Electrical and Electronics Engineers, New York, N.Y.

- [3] Recommandation UIT-R BS.1770-1, *Algorithmes de mesure de l'intensité sonore des programmes audio et des niveaux de crête vrais des signaux audio*, Union internationale des télécommunications, Genève, 2007.
- [4] ATSC, *Digital Audio Compression (AC-3, E-AC-3) Standard*, Doc. A/52B, Advanced Television Systems Committee, Washington, D.C., 14 juin 2005.
- [5] IEC 60268-17 (1990), *Équipements pour systèmes électroacoustiques. Partie 17 : Indicateurs de volume normalisés*.
- [6] SMPTE. *SMPTE RP 200 – Relative and Absolute Sound Pressure Levels for Motion-Picture Multichannel Sound Systems*, Society of Motion Picture and Television Engineers, White Plains, N.Y.
- [7] SMPTE. *SMPTE 302M-2007 Television – Mapping of AES3 Data into MPEG-2 Transport Stream*, Society of Motion Picture and Television Engineers, White Plains, NY, 2007.
- [8] SMPTE. *SMPTE 2020-1-2008 Format of Audio Metadata and Description of the Asynchronous Serial Bitstream Transport, SMPTE 2020-2-2008 Vertical Ancillary Data Mapping of Audio Metadata – Method A, SMPTE 2020-3-2008 Vertical Ancillary Data Mapping of Audio Metadata – Method B*, Society of Motion Picture and Television Engineers, White Plains, NY, 2008.
- [9] Recommandation UIT-R BS.1534-1, *Méthode d'évaluation subjective du niveau de qualité intermédiaire des systèmes de codage*, Union internationale des télécommunications, Genève, Suisse.
- [10] Soulodre, G.A., Lavoie, M.C. et Norcross, S.G. *The Subjective Loudness of Typical Program Material*, 115^e convention de l'audio Engineering Society, 2003.
- [11] Soulodre, G.A. *Evaluation of Objective Loudness Meters*, 116^e convention de l'Audio Engineering Society, Berlin, Preprint 6161.
- [12] Recommandation UIT-R BS.775-2, *Système de son stéréophonique multicanal avec ou sans image associée*, Union internationale des télécommunications, Genève, Suisse.
- [13] Soulodre, G.A., and Lavoie, M.C, *Stereo and Multichannel Loudness Perception and Metering*, 119^e convention de l'Audio Engineering Society, San Francisco, Preprint 6618.
- [14] AES. *TD-1001, Multichannel sound systems and operations*, Audio Engineering Society, New York, NY, www.aes.org/technical/documents/AESTD1001.pdf (téléchargement gratuit). Il est à noter qu'une exception a été faite dans l'édition 2001 de ce document : les niveaux d'écoute de référence de la section 7.4.1 du document ont été supprimés.
- [15] Recommandation UIT-R BS.1771, *Spécifications des appareils de mesure indiquant l'intensité sonore et les niveaux de crête vrais*, Union internationale des télécommunications, Genève, Suisse.
- [16] Eargle, J. M., *Audio Monitoring in Contemporary Post-Production Environments*, SMPTE Journal, Society of Motion Picture and Television Engineers, White Plains, NY, January, 2005.
- [17] Holman, Tomlinson, *Surround Sound Up and Running*, Focal Press, 2^e édition, 2007, (original ISBN-10 0240808290), ISBN-13 978-0240808291.
- [18] AES. *AES3-2003, Digital input-output interfacing - Serial transmission format for two-channel linearly represented digital audio data*, Audio Engineering Society, New York, N.Y.

- [19] Dolby. *Model DP569 Dolby Digital Multichannel Audio Encoder User's Manual*, Dolby Laboratories, Inc., San Francisco, CA.
- [20] Dolby. *Dolby DP570 Multichannel Audio Tool User's Manual*, Dolby Laboratories, Inc., San Francisco, CA.
- [21] AES.: *AES17-1998 (r2004), AES standard method for digital audio engineering - Measurement of digital audio equipment*, Audio Engineering Society, New York, N.Y.
- [22] SMPTE. *RDD 6-2008, Description and Guide to the Use of the Dolby® E Audio Metadata Serial Bitstream*, Society of Motion Picture and Television Engineers, White Plains, N.Y.
- [23] SMPTE. *SMPTE 337-2008, Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface*, Society of Motion Picture and Television Engineers, White Plains, N.Y.
- [24] ATSC. *A/79, Conversion of ATSC Signals for Distribution to NTSC Viewers*, Advanced Television Systems Committee, Washington, D.C., 12 décembre 2008.
- [25] CEA. *CEA-CEB11-A, NTSC/ATSC Loudness Matching*, Consumer Electronics Association, Arlington, VA, septembre 2009.
- [26] ANSI. *ANSI C16.5-1942, Standard Volume Indicator*, American National Standards Institute, Washington, D.C.
- [27] IEC. *Norme IEC 60268-17, Équipements pour systèmes électroacoustiques – Partie 17 : Indicateurs de volume normalisés*, Commission électrotechnique internationale, Genève, 1990.
- [28] IEC: *Norme IEC 60268-10/Iia, Iib, 1991, Équipements pour systèmes électroacoustiques Partie 10 : Appareils de mesure des crêtes de modulation*, Commission électrotechnique internationale, Genève, 1991.
- [29] ANSI. *ANSI S1.4-1983 (R 2001) avec modif. S1.4A-1995, Specification for Sound Level Meters*, American National Standards Institute, Washington, D.C.
- [30] *Commercial Advertisement Loudness Mitigation (CALM) Act*, Pub. L. 111-311, 47 U.S.C. § 621 (2011).

3 DÉFINITION DES TERMES

Les termes, abréviations et unités utilisés sont ceux de la pratique de l'Institute of Electrical and Electronics Engineers (IEEE), tels qu'ils sont exposés dans les normes publiées par l'Institut [2]. Si une abréviation n'est pas employée par l'IEEE ou si une pratique de l'industrie diffère de celle de l'IEEE, l'abréviation en question est décrite dans la section 3.3 du présent document.

3.1 Indication de conformité

Cette section définit les termes de conformité utilisés dans le présent document :

Vital – Ce mot qualifie une mesure devant être suivie à la lettre (aucun écart n'est autorisé).

Devrait – Ce mot indique qu'une mesure est privilégiée, mais non nécessairement exigée.

Ne devrait pas – Ce terme signifie que la possibilité ou la mesure est peu souhaitable, mais non interdite.

À titre d'aide supplémentaire, les recommandations essentielles de ce document sont accompagnées de l'icône .

3.2 Traitement des éléments syntaxiques

Ce document contient des références symboliques aux éléments de syntaxe employés dans les sous-systèmes de codage audio, vidéo et de transport. Ces références sont signalées par l'emploi d'une police de caractères distincte (p. ex. *restricted*), peuvent contenir le caractère de soulignement (p. ex. *sequence_end_code*) et peuvent se composer d'une chaîne de caractères ne correspondant pas à un mot français (ou anglais) (p. ex. *dynrng*).

3.3 Sigles et abréviations

Les abréviations et sigles suivants sont employés dans le document.

AC-3 – Norme de compression audio numérique (AC-3, E-AC-3), telle qu'elle est décrite dans le document ATSC A/52B [4]

ATSC – Advanced Television Systems Committee

dB – Décibel

dB FS – Décibel par rapport à une onde sinusoïdale pleine échelle (*Full Scale*) (suivant AES17 [21])

dB TP – Décibel du niveau de crête vrai (*True Peak*) par rapport à une mesure de niveau de crête vrai pleine échelle de 100% (suivant UIT-R BS.1770, annexe 2 [3])

DRC – Contrôle de gamme dynamique

GPI – Interface universelle

LKFS – Intensité sonore (*Loudness*) avec pondération K, par rapport à une pleine échelle nominale (*Full Scale*), mesurée au moyen d'un appareil qui met en œuvre l'algorithme spécifié par l'UIT-R BS.1770 [3]. Une unité LKFS est équivalente à un décibel.

UIT – Union internationale des télécommunications

MVPD – Distributeur multicanal de programmes audiovisuels (comprend les opérateurs de service de radiodiffusion directe par satellite (DBS), les câblodistributeurs locaux et les câblodistributeurs multiréseaux)

SPL – Niveau de pression acoustique (*Sound Pressure Level*) en décibels par rapport à $20 \mu\text{N/m}^2$

VU – Vumètre [5]

PPM – Modulomètre de crête (crête-mètre)

3.4 Termes

Les termes suivants sont employés dans le document.

Élément d'ancrage (*Anchor Element*) – Point ou élément de référence de la sonie (intensité sonore perçue) autour duquel les autres éléments du contenu sont équilibrés pour produire le mixage sonore final du contenu ou qui importe le plus au téléspectateur raisonnable quand il règle la commande de volume.

BS.1770 – Officiellement UIT-R BS.1770 [3]. Cette recommandation définit un algorithme qui permet d'obtenir une valeur numérique indiquant la sonie du contenu mesuré. Les sonomètres et les outils de mesure dans lesquels est mis en œuvre l'algorithme BS.1770 indiquent l'intensité sonore en « LKFS ».

Contenu (*Content*) – Matériel ou substance du produit distribué par un diffuseur.

Niveau de dialogue (*Dialog Level*) – L'intensité sonore de l'élément d'ancrage, exprimée en LKFS.¹

dialnorm – Paramètre de métadonnées AC-3, de valeur numérique égale à la valeur absolue du niveau de dialogue, transporté dans le flux binaire AC-3. C'est un code de 5 bits non signé qui indique l'écart négatif du niveau de dialogue moyen par rapport à 0 LKFS. Les valeurs valides vont de 1 à 31. La valeur 0 est réservée. Les valeurs 1 à 31 sont interprétées -1 à -31. Le décodeur applique une valeur de réduction de gain égale à la différence entre -31 et la valeur dialnorm.

Dolby E – Dolby E est une technique de réduction de débit de données audio conçu pour l'utilisation en contribution et en distribution, qui comporte aussi des métadonnées Dolby E.

Métadonnées Dolby E – Métadonnées qui sont multiplexées dans le flux binaire Dolby E. Chaque élément de métadonnées est classé dans le type « professionnel » ou « grand public ». Les métadonnées Dolby E sont expliquées dans le document SMPTE RDD 6 [22].

Profil DRC (*DRC profile*) – Ensemble de paramètres décrivant la façon dans les métadonnées de contrôle de la gamme dynamique sont calculées.

Dispositif de contrôle de gain dans les fichiers (*file-based scaling device*) – Dispositif servant à appliquer une correction de gain globale au contenu audio enregistré dans des fichiers.

Synchro trame (*framesync*) – Abréviation de « synchroniseurs de trame ».

Couchage de la bande sonore finale (*layback*) – Étape de postproduction où l'on réunit la bande sonore finale avec l'image, après le montage, le mixage ou le rehaussement du son final.

Contenu de longue durée (*long form content*) – Matériel ou substance de spectacle ou d'émission (programme). En règle générale, la durée dépasse environ deux ou trois minutes.

Sonie (*perceived loudness*) – Intensité sonore perçue; importance de l'effet physiologique produit quand un son stimule l'oreille.

Intensité sonore (*measured loudness*) – Grandeur d'un signal audio mesurée au moyen d'appareils mettant en œuvre l'algorithme défini dans la recommandation UIT-R BS.1770 [3]. Il s'agit d'une approximation de la sonie.

Niveau de mixage (*mixing level*) – Paramètre de métadonnées facultatif dans le flux binaire AC-3, qui permet d'indiquer l'étalonnage du niveau de pression acoustique absolu du studio de mixage qui a produit le contenu.

Opérateur (*Operator*) – Télédiffuseur, radiodiffuseur, opérateur de service DBS, câblodistributeur local ou câblodistributeur multiréseau (MSO) ou autre distributeur multicanal de programmes audiovisuels (MVPD).

Contenu de courte durée (*short form content*) – Matériel ou substance de messages publicitaires, commerciaux, promotionnels ou publics. Aussi appelé contenu « interstitiel ». Généralement, la durée est inférieure à deux ou trois minutes.

Intensité sonore cible (*Target Loudness*) – Valeur spécifiée pour l'Élément d'ancrage (c.-à-d. Niveau de dialogue), établie pour faciliter l'échange de contenu entre un fournisseur et un opérateur.

Crête vraie (*true peak*) – Valeur maximale de la forme d'onde du signal dans le domaine temporel continu, mesurée suivant les indications de la recommandation UIT-R BS.1770 [3].

¹ Le terme « Niveau de dialogue » est fondé sur l'utilisation répandue du dialogue comme ancrage pour le mixage du contenu; historiquement, on a convenu que le dialogue constituait l'Élément d'ancrage de la plupart des programmes.

L'unité de mesure est le dB TP (c.-à-d. décibels par rapport à 100 % de la crête vraie nominale).

4 LE SYSTÈME AUDIO MULTICANAL AC-3

Le système audio AC-3 de l'ATSC est conçu pour fournir une reproduction du contenu original (non traité) à la sortie du décodeur AC-3 dans un récepteur, normalisée à une intensité sonore uniforme. Il donne la possibilité aux radiodiffuseurs de laisser à chaque auditeur la liberté d'exercer un certain contrôle sur le degré de réduction de la gamme dynamique, au besoin, pour l'adapter au mieux à ses conditions d'écoute. La partie traitement de la gamme dynamique du système est décrite à la section 9, mais son fonctionnement implique la réception d'un contenu normalisé de manière appropriée.

Le paramètre de métadonnées *dialnorm* est transmis au décodeur AC-3 en même temps que les signaux audio codés. La valeur du paramètre *dialnorm* indique l'intensité sonore de l'Élément d'ancrage du contenu. La valeur *dialnorm* d'un programme d'un niveau sonore très élevé peut être 12 et celle d'un programme de niveau sonore doux, 27. La sortie du décodeur AC-3 est dotée d'un atténuateur qui applique l'atténuation appropriée pour normaliser l'intensité sonore du contenu à -31 LKFS.

Si le paramètre *dialnorm* correspond à l'intensité sonore générale du contenu, les auditeurs sont en mesure de régler leur commande de « volume » au niveau d'écoute (sonie) voulu et n'auront pas à la régler de nouveau chaque fois qu'il y a une pause publicitaire et au retour au programme. Si tous les radiodiffuseurs utilisent le système correctement, l'intensité sonore sera aussi uniforme d'une chaîne à l'autre.

La section 7 décrit trois méthodes d'utilisation des métadonnées audio : métadonnées statiques, prédéfinies et dynamiques. Ces trois méthodes assurent la livraison d'une intensité sonore uniforme aux auditeurs; le radiodiffuseur est libre d'employer la technique qui convient le mieux à ses pratiques. Quelle que soit l'approche choisie, le bon fonctionnement du système repose sur la transmission d'une valeur *dialnorm* qui représente bien le Niveau de dialogue du contenu, c'est-à-dire au départ sur la mesure précise de l'intensité sonore.

5 MESURE DE L'INTENSITÉ SONORE

L'intensité sonore étant un phénomène subjectif, c'est l'ouïe humaine qui en est le meilleur juge. Placés dans un environnement de mixage connu, tel que celui qui est décrit dans la section 10 de la présente RP, les mixeurs audio chevronnés se servant de leur ouïe peuvent produire un son (contenu sonore) de programme bien équilibré, d'une intensité sonore remarquablement uniforme. Si tous les programmes et tous les messages publicitaires étaient produits avec une intensité sonore uniforme et si l'intensité sonore du mixage était conservée durant les étapes de production, de distribution et tout au long de la chaîne de livraison, les auditeurs ne subiraient aucun changement contrariant de l'intensité sonore à l'intérieur des programmes et d'un programme à l'autre.

Dans la mesure des signaux audio, il y a deux paramètres d'intérêt clés : le niveau de crête vrai du signal et son intensité sonore. La mesure du niveau de crête vraie permet au mixeur de protéger le programme contre l'écrtage et la mesure de l'intensité sonore, de protéger l'auditeur contre les variations irritantes de sonie. Bien que le mixeur équilibre un mélange en se servant de ses oreilles, la mesure objective de l'intensité sonore aide à maintenir uniforme la sonie à l'intérieur d'un programme et d'un programme à un autre.

Les vumètres et les crêtes-mètres courants ne mesurent ni l'intensité sonore, ni le niveau de crête vrai du signal. Les caractéristiques de nombreux appareils de mesure « électroniques » courants disponibles ne sont pas connues, ce qui contribue à la situation d'incohérence et de confusion que l'on retrouve dans la pratique aujourd'hui.

Cette PR présente des directives qui, s'ils sont suivies, assureront l'uniformité de l'intensité sonore tout en éliminant les risques d'écèlement du signal. Les techniques de mesure recommandées sont fondées sur les mesures d'intensité sonore et de niveau de crête vrai définies dans la recommandation UIT-R BS.1770 – *Algorithmes de mesure de l'intensité sonore des programmes audio et des niveaux de crête vrais des signaux audio* [3]. Les détails des deux mesures sont présentés à l'annexe A.

5.1 Aperçu général de la méthode de mesure de l'intensité sonore UIT-R BS.1770

L'intensité sonore est mesurée en intégrant la puissance pondérée des signaux audio dans tous les canaux, sur toute la durée du contenu. La structure générale de l'algorithme est représentée à la figure 5.1.

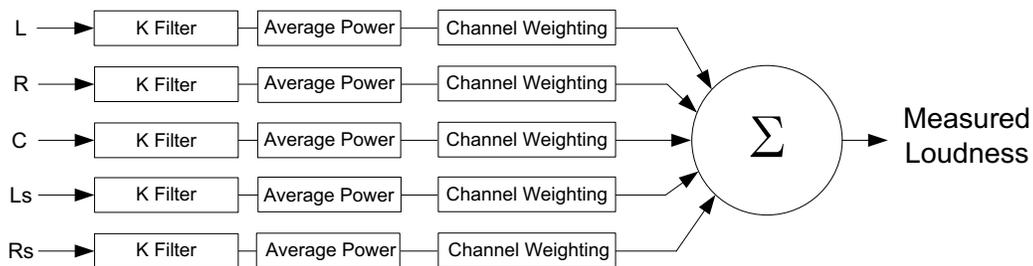


Figure 5.1 Algorithme de mesure de l'intensité sonore de l'UIT-R BS.1770²

K Filter	Filtre K
Average Power	Puissance moyenne
Channel Weighting	Pondération de canal
Measured Loudness	Intensité sonore
L	G
R	D
C	C
Ls	Ambiance gauche
Rs	Ambiance droite

La méthode BS.1770 a été validée au cours de tests d'écoute, en comparant les résultats qu'elle donnait à la sonie relative de séquences sonores mono, stéréo et multicanaux. L'intensité sonore est exprimée en LKFS. Une unité LKFS est identique à un décibel. Un programme de

² Le canal d'effets basses fréquences (LFE) n'est pas pris en compte dans l'algorithme de mesure courant de l'UIT-R BS.1770. Les utilisateurs de la présente PR devraient utiliser la version la plus récente de la recommandation UIT-R BS.1770.

-15 LKFS peut être réglé pour correspondre à la sonie d'un programme moins fort de -22 LKFS en l'affaiblissant de 7 dB.

L'intensité sonore de l'Élément d'ancrage (souvent le dialogue) du mixage est utilisée comme approximation de l'intensité générale du contenu. La mesure précise de l'intensité sonore de l'Élément d'ancrage (Niveau de dialogue) est nécessaire pour permettre aux opérateurs de distribuer le contenu aux auditeurs à des niveaux sonores uniformes.

Les détails complets des méthodes de mesure de l'intensité sonore et des niveaux de crête vrais sont présentés à l'annexe A.

5.2 Mesure

La norme de télévision numérique ATSC, A/53, exige la présence d'une valeur dûment établie de $dialnorm$ dans le flux élémentaire AC-3. Pour le contenu ayant des signaux audio s'étendant sur une plage de valeurs d'intensité sonore, l'objectif est que la valeur du paramètre $dialnorm$ corresponde à l'intensité sonore de l'Élément d'ancrage prédominant (généralement, le dialogue du programme).

Si le dialogue n'est pas l'Élément d'ancrage du contenu (p. ex. un programme musical), c'est l'intensité sonore de l'élément du mixage son sur lequel se fondera un téléspectateur raisonnable au moment de régler sa commande de volume qui devrait être mesurée et indiquée comme Niveau de dialogue du contenu.

Dans les cas où il n'est pas possible ou pratique de centrer la mesure sur l'Élément d'ancrage, une mesure intégrée sur une longue durée de l'intensité sonore de tous les éléments de la piste sonore devrait être exécutée.

Étant donné que la mesure de l'intensité sonore conformément à la recommandation BS.1770 est une mesure intégrée, les passages silencieux ont tendance à abaisser la valeur mesurée. Pour éviter cette situation, l'intégration peut être interrompue durant les segments de faible puissance. Certains organismes, dont l'UIT-R, étudient un procédé arrêt-pause-reprise automatique, généralement appelé « gating » (déclenchement périodique), que l'UIT-R prévoit ajouter ultérieurement à la recommandation BS.1770. Certains appareils peuvent offrir une fonction de déclenchement périodique. Toutefois, il n'existe pas encore de norme visant la mesure d'intensité sonore avec fonction de déclenchement périodique. Les utilisateurs devraient employer la version courante de l'UIT-R BS.1770 pour les mesures.

5.2.1 Mesure durant la production ou la postproduction

Durant le processus de production ou de postproduction, il est généralement possible d'identifier l'Élément d'ancrage (généralement le dialogue) du mixage son du programme et d'en mesurer l'intensité sonore avec précision. Un échantillon représentatif de l'ancrage devrait être mesuré conformément à la recommandation BS.1770 [3]. S'il est impossible d'identifier et de mesurer seul l'Élément d'ancrage, on devrait alors mesurer l'intensité sonore de longue durée de tous les éléments de la piste sonore, sur toute la durée, et la signaler comme Niveau de dialogue.

5.2.2 Mesure en temps réel en cours de production (événement en direct)

Le but des mesures d'intensité sonore durant un événement en direct est de guider le mixeur de son pour produire un contenu ayant l'Intensité sonore cible ou préétablie. Le principe de la mesure de l'intensité sonore de l'Élément d'ancrage du contenu s'applique aussi aux productions en direct, mais la mesure est faite en temps réel, à mesure que progresse la production.

Certains sonomètres BS.1770 présentent un mode de mesure dite à fenêtre glissante ou à moyenne mobile indiquant l'intensité sonore d'un segment antérieur (en général de 3 à

10 secondes) du programme. D'autres appareils peuvent utiliser une certaine représentation graphique de l'intensité sonore fondée sur la recommandation BS.1770 en complément d'une simple valeur numérique. Ces modes de mesure devraient être employés pour aider le mixeur de son à obtenir l'Intensité sonore cible. Ces mesures sont particulièrement utiles pour le mixage en milieu bruyant ou quand il est impossible de maintenir un niveau de contrôle uniforme.

La mesure à fenêtre glissante est perturbée lorsque les périodes de silence constituent une partie notable de la durée de la fenêtre, c'est pourquoi il convient d'employer cette technique avec prudence.

La mesure continue de l'intensité sonore du début de la production jusqu'à un point ultérieur quelconque peut aussi fournir une indication de l'intensité sonore moyenne de la production.

5.2.3 Mesure de contenu de longue durée fini

Il peut être difficile d'identifier l'Élément d'ancrage d'un programme fini et d'en mesurer l'intensité sonore. Dans la mesure du possible, une partie du contenu représentative de l'Élément d'ancrage devrait être isolée et soumise aux mesures. En l'absence d'un Élément d'ancrage précis, on devrait mesurer l'intensité sonore de l'élément de contenu sur lequel se fonderait un téléspectateur raisonnable pour régler le volume. Si aucune de ces méthodes n'est réalisable ou pratique, l'intensité sonore de tous les éléments du contenu devrait être mesurée.

5.2.4 Mesure de contenu de courte durée

Les mixeurs de son de contenu de courte durée devraient suivre la pratique définie pour la production/postproduction, telle qu'elle est décrite dans la section 5.2.1 ci-dessus. En règle générale, toutefois, les opérateurs n'ont pas accès au contenu de courte durée durant la production et l'Élément d'ancrage est rarement isolé. Il est donc plus pratique de mesurer l'intensité sonore de tous les éléments de la piste sonore sur toute la durée du contenu de courte durée.

5.2.5 Mesure fondée sur des fichiers

Les méthodes de mesure décrites ci-dessus s'appliquent aussi au contenu existant dans des fichiers numériques. En outre, le stockage dans des fichiers facilite l'automatisation des mesures d'intensité sonore et l'examen de la valeur dialnorm (si le paramètre est défini) qui a pu être attribuée au contenu. L'intensité sonore du contenu peut être ajustée au besoin à la valeur de l'Intensité sonore cible ou désirée, en appliquant une correction de gain globale³, ou la valeur dialnorm peut être réécrite pour la faire correspondre à l'intensité sonore mesurée du contenu.

6 NIVEAU SONORE CIBLE ET NIVEAU DE CRÊTE VRAI POUR LA DISTRIBUTION OU L'ÉCHANGE DE CONTENU

Pour la distribution et l'échange de contenu sans métadonnées (et lorsqu'il n'existe pas de disposition préalable prise par les intéressés visant l'intensité sonore), l'Intensité sonore cible devrait être de -24 LKFS. De faibles écarts de mesure par rapport à cette valeur, allant jusqu'à environ ± 2 dB, sont prévus en raison de l'incertitude des mesures et sont acceptables (en raison de la zone de confort – voir l'annexe E). L'intensité sonore du contenu *ne devrait pas* être ajustée délibérément à la limite inférieure ou supérieure de cette plage. Le niveau de crête vrai devrait être maintenu sous les -2 dB TP afin de laisser une marge de sécurité pour éliminer les risques d'écèlement en raison du traitement aval (comme le codage audio utilisé pour la distribution).

³ Si le contenu est stocké dans un format à débit binaire réduit (codé), il peut être nécessaire de décoder le contenu, d'en modifier le niveau, puis de le recoder.

7 POINTS À CONSIDÉRER DANS LA GESTION DES MÉTADONNÉES AYANT UN EFFET SUR L'INTENSITÉ SONORE DES PROGRAMMES AUDIO

L'encodeur AC-3 permet de définir 28 paramètres de métadonnées portant sur les caractéristiques des signaux audio présents dans le flux binaire (voir l'annexe G). Ces paramètres peuvent être classés en trois groupes :

Métadonnées informatives, groupe composé de sept paramètres facultatifs pouvant être employés pour décrire le son codé. Ces paramètres n'ont pas d'effet sur le codage ou sur l'expérience d'écoute du son décodé à la maison.

Métadonnées de contrôle de base, groupe composé de 19 paramètres déterminant la compression de la gamme dynamique, le mixage réducteur, le décodage de matrice et le filtrage utilisés dans certains modes de fonctionnement de l'encodeur professionnel et du décodeur grand public. L'optimisation du réglage de ces paramètres pour chaque programme peut améliorer l'expérience d'écoute dans diverses conditions d'écoute et pour certains types de contenu. Toutefois, les valeurs par défaut peuvent être employées sans nuire à l'expérience d'écoute.

Métadonnées de contrôle essentielles, groupe composé de deux paramètres qui sont cruciaux pour assurer l'encodage et le décodage appropriés :

- **Paramètre de mode canal** (*acmod*), qui devrait être choisi correctement pour lancer le formatage de canaux approprié dans le décodeur, correspondant au contenu. La valeur erronée de ce paramètre peut modifier une transmission et entraîner la perte de dialogue au cours du codage d'un programme 5.1; p. ex. codage d'une piste sonore 5.1 canaux avec la valeur de métadonnée 2/0.
- **Paramètre de Niveau de dialogue** (*dialnorm*), que la norme A/53 exige de dûment régler pour éliminer les risques de variation de sonie (potentiellement grave) durant les transitions de contenu sur une chaîne ou dans le changement de chaîne de TVN. Les valeurs erronées de *dialnorm* peuvent conduire à une variation d'intensité sonore atteignant 30 dB.

Le paramètre *dialnorm* mis à part, les valeurs par défaut peuvent être utilisées pour la plupart des autres paramètres de métadonnées avec des résultats acceptables. Une fois que les mixeurs et les producteurs connaissent bien ces paramètres grâce à un monitoring avec des systèmes d'émulation disponibles, ils peuvent choisir des valeurs qui optimisent davantage la présentation du contenu de leurs programmes.

7.1 Importance de *dialnorm*

Le transport et le bon réglage de la valeur de *dialnorm* sont obligatoires en télédiffusion numérique aux États-Unis; voir le document ATSC A/53, Part 5:2010, Section 5.5, *Dialogue Level* [1].

La présente PR indique les méthodes permettant d'assurer l'uniformité de l'intensité sonore en télévision numérique grâce à l'utilisation appropriée de la métadonnée *dialnorm* pour tous les contenus, et donc de se conformer à la norme A/53. Un grand nombre des principes pour la gestion efficace de *dialnorm* s'appliquent aussi à la gestion d'autres paramètres de métadonnées AC-3.

Comme l'indique la section 6, de légers écarts de mesure pouvant atteindre environ ± 2 dB sont prévus, et ils peuvent entraîner de légères variations entre la valeur *dialnorm* et l'intensité sonore réelle du programme. Ces faibles variations sont acceptables (en raison de la zone de confort – voir l'annexe E); toutefois, les opérateurs *ne devraient pas* diffuser intentionnellement à la limite inférieure ou supérieure de cette plage.

7.2 Modes de gestion des métadonnées

L'exigence visant l'exactitude des valeurs de dialnorm, d'acmod (mode canal) et d'autres métadonnées peut être remplie de trois façons différentes, à la discrétion de l'opérateur :

- **Métadonnées statiques** : le Niveau de dialogue de l'encodeur AC-3 est réglé à une valeur « fixe » et les Niveaux de dialogue de contenu sont alignés sur cette valeur.
- **Métadonnées prédéfinies** : des « préréglages » sont programmés dans l'encodeur AC-3, chacun ayant une valeur dialnorm différente et la valeur est appliquée par l'intermédiaire d'une interface universelle (GPI) ou autre interface de contrôle.
- **Métadonnées dynamiques** : l'encodeur AC-3 est configuré pour recevoir des métadonnées externes. Un système de métadonnées dialnorm « agile » en amont peut être utilisé pour fournir des valeurs dialnorm dynamiques à l'encodeur, en fonction des variations de l'intensité sonore aux frontières des contenus.

Quand elles sont appliquées judicieusement, les trois méthodes fournissent un résultat final conforme et acceptable pour le consommateur. La majorité de l'explication de la présente section de la PR est centrée sur le paramètre dialnorm. Le lecteur est invité à consulter l'Annexe G et à rechercher les renseignements sur la façon dont les autres paramètres de métadonnées peuvent influencer sur le codage.

L'opérateur a aussi la possibilité d'appliquer une méthode hybride en choisissant une méthode pour la gestion de l'intensité sonore et une autre pour les autres métadonnées, p. ex. utiliser une valeur dialnorm statique, mais changer de mode de canal au besoin.

7.3 Utilisation de métadonnées dialnorm statiques

Le concept de valeur dialnorm statique consiste simplement à attribuer une valeur « fixe » au paramètre dialnorm de l'encodeur AC-3 pour le réseau ou le système de radiodiffusion et à aligner le signal d'entrée audio de l'encodeur sur cette valeur. L'opérateur peut choisir n'importe quelle valeur de dialnorm de 31 à 1; toutefois, pour être conforme à la norme A/53, l'opérateur doit employer une valeur égale à l'intensité sonore moyenne du dialogue de tout le contenu. Voir la figure 7.1.

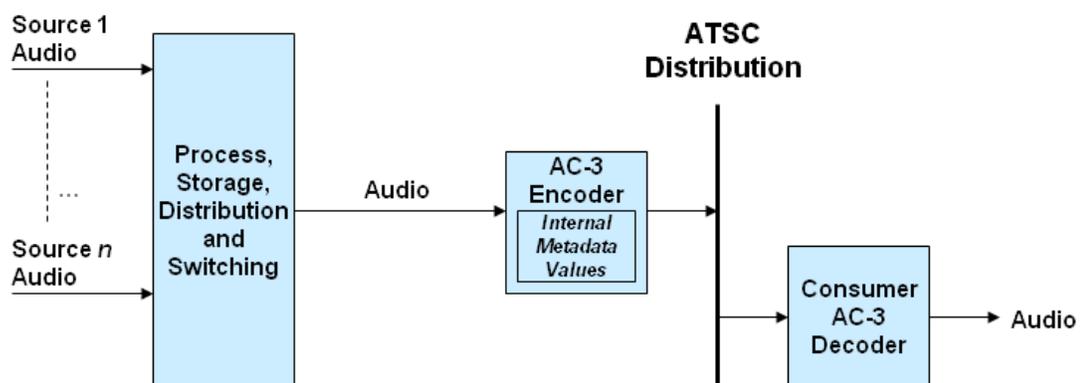


Figure 7.1 Concept de métadonnées statiques

Source 1 Audio	Audio source 1
Source n Audio	Audio source n
Process, Storage, Distribution and Switching	Traitement, stockage, distribution et commutation
Audio	Audio
AC-3 Encoder <i>Internal Metadata Values</i>	Encodeur AC-3 <i>Valeurs de métadonnées internes</i>
ATSC Distribution	Distribution ATSC
Consumer AC-3Decoder	Décodeur AC-3 du consommateur
Audio	Audio

7.3.1 Établissement de la valeur dialnorm par la méthode de la moyenne longue durée

L'opérateur peut obtenir une première approximation de conformité à A/53 en mesurant l'intensité sonore moyenne du matériel de longue durée à la sortie de la station et en attribuant cette valeur au paramètre dialnorm de l'encodeur AC-3. La période pour l'établissement de la moyenne doit être choisie de manière à inclure tous les types de contenu. Si le Niveau de dialogue de segments particuliers de contenu s'écarte considérablement de cette moyenne longue durée, le paramètre dialnorm ne correspondra pas au Niveau de dialogue de ce contenu. Cette situation doit être traitée et corrigée par le créateur du programme ou par l'opérateur (voir les sections 7.3.2 et 7.3.3.). Cette méthode peut ne pas s'appliquer aux opérateurs utilisant un contenu dont la gamme dynamique est intentionnellement étendue.

7.3.2 Établissement de la valeur dialnorm pour la production

Les spécifications de distribution de contenu devraient préciser l'Intensité sonore cible de tous les contenus. Elles établissent l'ancrage pour superposer la musique et les effets sur la piste sonore. L'intensité sonore devrait être mesurée au moyen d'un sonomètre, suivant la recommandation UIT-R BS.1770 [3], pour confirmer l'intensité sonore moyenne du dialogue. Le fournisseur devrait indiquer l'intensité sonore moyenne réelle avec le produit livrable. Dans la mise en œuvre de cette pratique, la collaboration entre le fournisseur de contenu et le destinataire est nécessaire pour assurer la bonne gestion de l'intensité sonore.

7.3.3 Contenu non conforme à l'Intensité sonore cible

Si l'opérateur a besoin d'utiliser un contenu non conforme à la valeur de l'Intensité sonore cible établie, un gain ou une perte de compensation doit être inséré. Si la différence est inconnue, il faut mesurer l'intensité sonore du contenu avant d'appliquer la compensation.

7.3.4 Valeur dialnorm et contrôle de qualité d'intensité sonore

Pour assurer la correspondance entre la valeur dialnorm et l'intensité sonore, l'opérateur doit avoir recours à un sonomètre durant le contrôle de qualité et, au besoin, apporter les ajustements nécessaires pour faire en sorte que l'intensité sonore corresponde à la valeur cible.

7.3.5 Réglage dialnorm d'émission pour la conformité à la norme A/53

Un opérateur recevant un contenu d'intensité sonore fixe et qui ne fera l'objet d'aucun traitement ou ajustement de gain après le récepteur, devrait régler la valeur de dialnorm dans l'encodeur AC-3 d'émission à la valeur de l'intensité sonore spécifiée par le créateur du contenu (souvent indiquée dans un contrat, un document de spécification des signaux, etc.). Si un gain ou une perte

fixe est appliquée dans la chaîne de signaux, la valeur dialnorm de l'encodeur AC-3 devrait être compensée en conséquence de l'intensité sonore indiquée par le créateur.

Par exemple, si le créateur livre un contenu audio ayant une intensité sonore de -24 LKFS et qu'aucun gain ou perte n'est produit dans la chaîne, la valeur dialnorm serait fixée à 24⁴. Cependant, si un gain de 3 dB est ajouté, ramenant l'intensité sonore à -21 LKFS, la valeur de dialnorm serait fixée à 21. Si c'est plutôt une perte de 2 dB qui est introduite, l'intensité sonore passerait à -26 LKFS, et la valeur de dialnorm serait de 26.

Si un processeur d'intensité sonore est employé, la valeur dialnorm de l'encodeur AC-3 devrait être fixée à la valeur du Niveau de dialogue à la sortie du processeur audio.

7.3.6 Avantage du paramètre dialnorm statique

Un système avec dialnorm fixe présente un risque minimal pour le contenu. Le paramètre dialnorm statique a l'avantage d'être simple, sans nécessiter la gestion de données ou de matériel de métadonnées supplémentaire. Cette approche peut être utilisée avec tous les encodeurs AC-3 et elle est la seule possible quand on utilise un encodeur sans entrée de métadonnées ou commande GPI externe.

7.4 Utilisation de métadonnées dialnorm prédéfinies

Si l'opérateur doit composer avec un petit nombre de modifications discrètes de la valeur dialnorm ou d'autres paramètres de métadonnées, certains systèmes d'encodage AC-3 peuvent être configurés pour permettre la sélection d'une valeur de métadonnées dans une liste prédéfinie, par l'intermédiaire d'une commande externe; p. ex. au moyen d'une fermeture de contact avec une GPI. Cette méthode exige des déclencheurs externes de GPI pour la signalisation de valeurs prédéfinies précises à partir de la liste de lecture ou du commutateur d'automatisation. Elle est souvent employée pour changer de mode de codage, de stéréo à 5.1, même lorsque dialnorm est statique et ne prend qu'une seule valeur. Voir la figure 7.2.

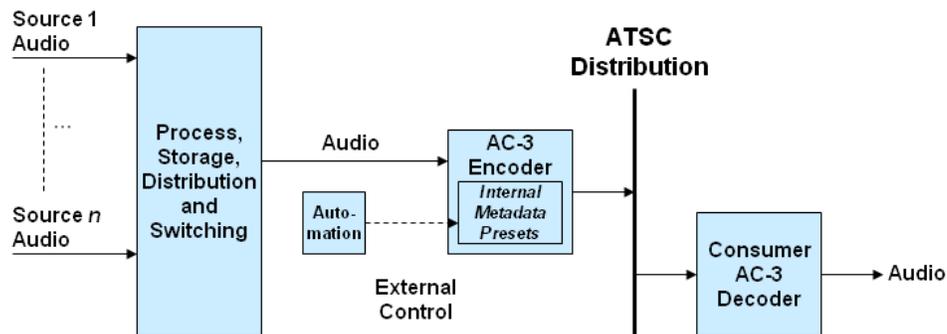


Figure 7.2 Concept de métadonnées prédéfinies

Source 1 Audio	Audio source 1
Source <i>n</i> Audio	Audio source <i>n</i>
Process, Storage, Distribution and Switching	Traitement, stockage, distribution et commutation

⁴ Comme il est signalé dans la section 3.4, dialnorm est exprimé comme un entier non signé.

Audio	Audio
Automation	Automatisme
External Control	Commande externe
AC-3 Encoder <i>Internal Metadata Presets</i>	Encoder AC-3 <i>Valeurs de métadonnées internes prédéfinies</i>
ATSC Distribution	Distribution ATSC
Consumer AC-3 Decoder	Décodeur AC-3 de consommateur
Audio	Audio

7.4.1 Mise en œuvre

La mise en œuvre de métadonnées prédéfinies est semblable à celle des métadonnées statiques. Les valeurs préétablies sont chargées dans l'encodeur AC-3 en fonction des différences connues d'intensité sonore du contenu. La conformité à la norme A/53 exige que l'intensité sonore du contenu livré corresponde à l'une des valeurs prédéfinies et que le système d'automatisation soit programmé pour changer de valeur prédéfinie en fonction de différentes valeurs d'intensité sonore de contenu et de mode de canal.

7.4.2 Exigence de synchro trame AC-3

Certains encodeurs AC-3 réinitialisent et perturbent la sortie du flux binaire audio quand un préréglage est changé. Selon le type d'encodeur ATSC utilisé, cette modification peut produire un bruit parasite audible en ondes. Pour éliminer ce risque, il peut être nécessaire de prévoir un synchro trame AC-3 à la sortie de l'encodeur AC-3 pour stabiliser la source AC-3 de l'encodeur ATSC.

7.5 Utilisation de métadonnées dialnorm dynamiques

Un système de métadonnées « agile » permet d'attribuer différentes valeurs dialnorm à différents contenus ayant une intensité sonore différente. L'opération consiste à incorporer le paramètre dialnorm dans le flux binaire de métadonnées accompagnant le contenu dans un emplacement en « amont ». Les métadonnées sont extraites juste avant l'encodeur AC-3 et raccordées à une entrée de métadonnées série externe. La valeur dialnorm change de façon appropriée aux frontières des contenus. Voir la figure 7.3.

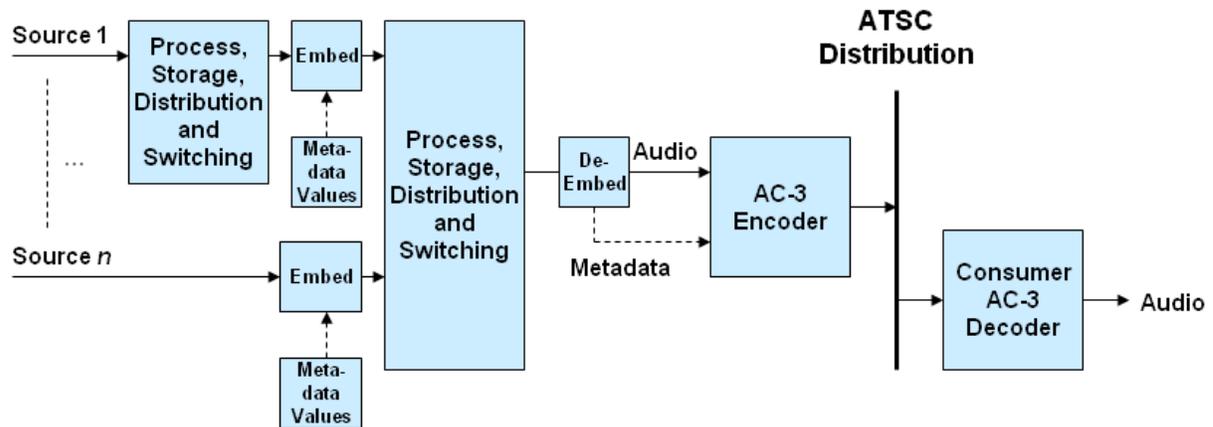


Figure 7.3 Concept de métadonnées dynamiques

Source 1	Source 1
Source n	Source n
Process, Storage, distribution and Switching	Traitement, stockage, distribution et commutation
Embed	Incorpore
Metadata Values	Valeurs de métadonnées
De-Embed	Extrait
Metadata	Métadonnées
Audio	Audio
AC-3 Encoder	Encodeur AC-3
ATSC Distribution	Distribution ATSC
Consumer AC-3 Decoder	Décodeur AC-3 du consommateur

7.5.1 Déploiement du système

Quand un opérateur de réseau utilise la méthode d'attribution dynamique des métadonnées, la même méthode doit être adoptée dans toutes les installations des stations de radiodiffusion ou de la tête de réseau du MVPD qui reçoit le contenu du réseau. Cette technique nécessite le déploiement de matériel d'encodage et de décodage complexe à tous les points d'entrée, de sortie, de contrôle et de traitement de la chaîne de distribution, et ce, du point d'origine des métadonnées jusqu'à tous les encodeurs AC-3. Il est essentiel que les métadonnées dynamiques atteignent l'encodeur AC-3. Plusieurs méthodes de distribution et de stockage de métadonnées dynamiques sont disponibles, qui peuvent être utilisées de manière autonome ou combinée — les métadonnées AC-3 (comme il est décrit dans l'annexe G) constituent un sous-ensemble grand public des métadonnées Dolby E décrites dans le document SMPTE RDD 6 [22]. Ces données peuvent être transportées sur des liaisons de données série, comme données auxiliaires verticales (VANC) ou comme données transportées dans les flux binaires compressés. Elles peuvent aussi être stockées dans des systèmes de fichiers.

7.5.1.1 Métadonnées Dolby E sur liaison série

Les métadonnées Dolby E dans leur forme bande de base peuvent être transportées par liaison série. Ce procédé nécessite une couche série réservée demeurant étroitement synchronisée dans le temps avec les signaux audio et vidéo.

7.5.1.2 Métadonnées Dolby E en VANC

Les métadonnées Dolby E peuvent aussi être incorporées dans les VANC des systèmes numériques série de définition standard ou haute définition, puis extraites en aval suivant les indications de la norme SMPTE 2020 [8]. Cette méthode peut nécessiter des multiplexeurs et des démultiplexeurs et exige la prise en charge par le matériel de stockage vidéo, de codage, de traitement et de distribution et la capacité de passer le signal VANC en le laissant intact. Certains dispositifs de stockage ont une capacité VANC limitée ou nulle.

7.5.1.3 Métadonnées et codecs

Certains systèmes utilisés pour les applications de liaison secondaire, de distribution et de stockage ont aussi la capacité de transporter les métadonnées Dolby E. Ces systèmes incluent le flux binaire Dolby E compressé lui-même et des formats propriétaires⁵ qui nécessitent l'emploi d'encodeurs et de décodeurs audio spécialisés. Ils peuvent aussi nécessiter du matériel capable de décaler la synchronisation vidéo pour compenser la latence d'encodage et de décodage introduite. La plupart des équipements de vidéo numérique professionnels peuvent être configurés pour passer ces signaux codés dans des canaux audio numériques standard, conformes à SMPTE 337 [23].

7.5.1.4 Métadonnées enregistrées dans des fichiers

Il existe un grand nombre de techniques de stockage de métadonnées Dolby E dans des fichiers, certaines étant normalisées et quelques-unes, exclusives. Ce sujet est hors du propos du présent document.

7.5.2 Technique de production en direct

Dans la production en direct avec la méthode des métadonnées dynamiques, l'ingénieur du son au mixage de la production télévisuelle sélectionne une intensité sonore cible particulière, mais arbitraire, pour chaque programme, en tenant compte de la dynamique, de la marge de sécurité et du type d'ambiance du programme. Ce paramètre établit l'ancrage de l'intensité sonore pour la superposition de la musique et des effets sur la piste sonore. Suivant la spécification du produit livrable, l'Intensité sonore cible est soit transportée dans le signal codé par la valeur du paramètre dialnorm, soit communiquée au distributeur pour codage ultérieur.

7.5.3 Technique de production en différé

Dans la postproduction, avec la méthode des métadonnées dynamiques, l'intensité sonore du mixage final est déterminée soit durant la production du programme ou une fois la production terminée. Suivant la spécification du produit livrable, l'Intensité sonore cible est soit transportée dans le signal codé par la valeur du paramètre dialnorm, soit communiquée au distributeur pour codage ultérieur. L'insertion dans le produit livrable peut être exécutée dans un processus de réenregistrement ou en ayant recours à la fonction de prélecture offerte dans certains magnétoscopes. Voir la figure 7.4.

⁵ P. ex. *e-squared* de Linear Acoustic et *Quad Phase Aligned* de Tandberg.

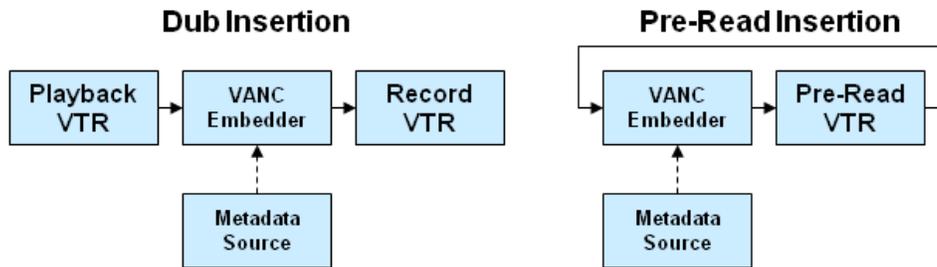


Figure 7.4 Options d'insertion des métadonnées

Dub Insertion	Insertion en cours de réenregistrement
Playback VTR	MagnétoSCOPE de lecture
VANC Embedder	Incorporateur de VANC
Metadata Source	Source de métadonnées
Record VTR	MagnétoSCOPE d'enregistrement
Pre-Read Insertion	Insertion en prélecture
Pre-Read VTR	MagnétoSCOPE de prélecture

7.5.4 Contrôle de la production

La piste sonore doit faire l'objet de mesures, conformément à la recommandation UIT-R BS.1770 [3], pour confirmer que l'intensité sonore moyenne pour toute la durée de la production correspond à la valeur dialnorm choisie. (Voir la section 5.)

7.5.5 Métadonnées semi-dynamiques

Un opérateur peut employer un système de métadonnées dynamiques, mais choisir de simplifier la création et l'insertion des métadonnées en précisant les valeurs d'Intensité sonore cible statiques que devront utiliser les fournisseurs de contenu.

7.5.6 Incidence de la perte de métadonnées sur le contenu

Un risque lié à l'utilisation d'un système de métadonnées dynamiques est la possibilité qu'il existe un écart élevé d'intensité sonore d'un programme à un autre et entre les chaînes en cas de perte des métadonnées. Tous les encodeurs AC-3 ayant une entrée de métadonnées externe présentent une fonction de retour aux paramètres définis, dite de « réversion », destinée à atténuer l'incidence de la perte de métadonnées. Doté de cette fonction, l'encodeur peut être configuré pour conserver la valeur de métadonnée la plus récente ou revenir aux valeurs prédéfinies par l'opérateur. Bien que la réversion puisse réduire au minimum les effets sur le consommateur, l'erreur dans les paramètres d'intensité sonore ou d'autres métadonnées (comme le mode de canal) peut tout de même être grave. Le paramètre de réversion devrait être choisi de manière à réduire au minimum l'incidence de la perte de métadonnées sur le contenu présenté.

7.5.7 Hybride statique-dynamique

Dans certains cas, un opérateur peut choisir d'utiliser intentionnellement la fonction de réversion pour traiter un contenu sans métadonnées.

Il est crucial que les opérateurs choisissent les réglages appropriés pour tous les paramètres de métadonnées prédéfinis pour la fonction de réversion et qu'ils veillent particulièrement à ce

que l'intensité sonore du contenu distribué sans métadonnées corresponde au paramètre dialnorm prédéfini pour la réversion. La fonction de réversion peut aussi servir de protection contre la perte de métadonnées, même si les paramètres de métadonnées de la réversion ne coïncident pas exactement avec ceux du contenu. En cas de perte de métadonnées, tout le contenu codé après la réversion est soumis à ces paramètres. Il est particulièrement crucial que le mode de canal soit réglé de manière à protéger tout le contenu dans n'importe quelle situation. L'utilisation accidentelle du mode de canaux 2/0 pour un contenu 5.1 aurait pour effet d'éliminer les canaux 3 à 6 de l'audio codé et de compromettre le contenu.

7.5.8 Avantages des métadonnées dynamiques

Le système dynamique présente la plus grande souplesse pour le fournisseur de contenu sans imposer de limites à la créativité.

8 MÉTHODES DE CONTRÔLE EFFICACE DE L'INTENSITÉ SONORE ENTRE PROGRAMME ET CONTENU INTERSTITIEL

En raison de sa gamme dynamique étendue et des nouvelles techniques de gestion de l'intensité sonore, le système audio de télévision numérique ATSC (AC-3) présente la possibilité de changements d'intensité sonore indésirables dans les transitions de contenus (passage d'un élément de contenu à un autre) si le système n'est pas géré de façon appropriée. Cette situation est une cause de contrariété connue des téléspectateurs, qui les force souvent à régler le volume de leur appareil durant les transitions pour maintenir l'intensité sonore à un niveau confortable. L'application d'une bonne gestion de l'intensité sonore des programmes de TVN permet de remédier à ce problème.

Le système AC-3 intègre la technologie nécessaire pour atténuer les variations d'intensité sonore se produisant au passage des programmes aux contenus interstitiels. Les techniques prévues sont décrites ci-dessous.

8.1 Solutions efficaces

La variation importante d'intensité sonore durant les transitions peut être contrôlée efficacement en veillant à ce que le paramètre dialnorm corresponde bien au Niveau de dialogue de tout le contenu :

8.1.1 Pour les opérateurs employant un système à dialnorm statique (voir la section **Error! Reference source not found.**)

- a) Veiller à ce que tout le contenu réponde à l'Intensité sonore cible et que l'intensité sonore longue durée corresponde à la valeur dialnorm.
- b) Employer un dispositif de contrôle de gain de fichiers pour faire correspondre à la valeur cible l'intensité sonore longue durée du contenu de fichiers non conformes.
- c) Employer un dispositif de traitement de l'intensité sonore en temps réel pour faire correspondre à la valeur cible le contenu en temps réel non conforme.

8.1.2 Pour les opérateurs employant un système dialnorm dynamique (voir la section 7.5)

- a) Veiller à ce qu'au cours de la production, de la postproduction ou de l'ingestion, l'intensité sonore du contenu soit mesurée (voir section 5.2) et que le contenu soit étiqueté de la bonne valeur dialnorm, correspondant à l'intensité sonore réelle du contenu particulier.

- b) Employer un dispositif de mesure et de création média de fichiers pour régler dialnorm à l'intensité sonore moyenne du contenu particulier.
- c) Employer un dispositif de traitement en temps réel servant à faire correspondre le contenu à une intensité sonore précise. Appliquer une valeur dialnorm correspondant à l'intensité sonore de tout le contenu traité par ce dispositif.

8.2 Conditions hostiles

Conditions notables pouvant nuire aux transitions entre les programmes et le contenu interstitiel aux extrémités des contenus :

- Les fournisseurs de contenu augmentent souvent l'effet dramatique en utilisant la dynamique de programme et en manipulant l'intensité sonore pour obtenir un effet voulu sur les auditeurs. Cette opération est parfois exécutée à la fin des segments de programmes précédant une pause publicitaire.
- Une variation extrême, allant au-delà des limites de la zone de confort (voir l'annexe E), peut nécessiter qu'un auditeur règle le volume pour compenser le changement important, temporaire, de l'intensité sonore. Quand un message publicitaire ou d'autopublicité est présenté à l'entrée ou à la sortie des pauses, l'auditeur peut devoir régler de nouveau le volume pour l'écoute du contenu de courte durée. Cette situation s'est avérée contrariante pour les téléspectateurs.

8.3 Recommandations sommaires

Voici les recommandations visant à diminuer l'incidence négative des variations d'intensité sonore au passage des programmes aux contenus interstitiels :

- 1) Dans un système à métadonnées statiques, veiller à bien cibler l'intensité sonore moyenne, ou
- 2) Dans un système à métadonnées dynamiques, veiller à ce que la valeur dialnorm de création média corresponde à l'intensité sonore mesurée du contenu.
- 3) Sensibiliser le fournisseur de contenu à l'utilisation potentiellement abusive de la gamme dynamique et à la possibilité que les téléspectateurs se plaignent en cas de variation importante d'intensité sonore.
 - a) Expliquer cette situation dans les spécifications de distribution de contenu que l'opérateur exige du fournisseur :
 - i) Décrire l'incidence négative que produit sur les téléspectateurs le mixage dépassant les limites de la zone de confort d'écoute (voir l'annexe E) au passage et au retour des pauses.
 - ii) Décrire l'intensité sonore prévue du matériel interstitiel à l'intention du fournisseur de programmes, dans un effort de sensibilisation à cette situation, à l'insatisfaction possible de l'auditeur et aux répercussions négatives potentielles sur la présentation. □
 - b) Veiller à utiliser le niveau de pression acoustique (SPL) approprié dans les salles utilisées pour le mixage et le contrôle des pistes sonores de contenu.
 - i) Renvoyer les fournisseurs de contenu à la section 10 de la présente PR, appuyant l'utilisation d'un niveau de contrôle SPL convenable durant la postproduction de contenu. Le niveau de contrôle SPL choisi devrait être approprié pour les dimensions de la salle de mixage et tenir compte de l'environnement d'écoute du téléspectateur type de TVN. Un niveau bas de contrôle SPL en salle de contrôle postproduction

- entraîne des mixages dont l'intensité sonore sera plus élevée et dont la dynamique sera plus contenue et appropriée, que les environnements sonores classiques, plus « forts ». Il est essentiel de bien choisir l'environnement de contrôle de postproduction ou de reproduction pour établir des niveaux de mixage appropriés pour la télévision numérique.
- ii) Songer à spécifier une valeur de crête vraie maximale pour la piste sonore. Cette pratique limite la gamme dynamique en réduisant la marge de sécurité. Elle permet aux téléspectateurs de régler le volume général avec un risque réduit de variation importante d'intensité sonore.
 - 4) On ne devrait pas compter sur le système de DRC AC-3 pour contrôler les variations d'intensité sonore au passage des programmes au contenu interstitiel.

8.4 Insertion de publicité locale par la station de télévision et le MVPD

Dans le cas d'insertion de messages publicitaires ou de segments locaux par la station de télévision ou le MVPD, l'opérateur devrait veiller à ce que le Niveau de dialogue de l'insertion locale coïncide avec le réglage dialnorm du flux audio inséré.

Si le produit du créateur destiné au réseau est décodé en bande de base, il faut mesurer l'intensité sonore du flux audio décodé et régler la valeur dialnorm du ré-encodeur AC-3 pour qu'elle corresponde à l'intensité sonore mesurée, pour l'étape de codage suivante. Dans ce cas, soit que l'opérateur modifie l'intensité sonore du contenu du créateur pour qu'elle corresponde à la valeur cible du système de l'opérateur, soit qu'il utilise la valeur d'intensité sonore du contenu du créateur (mesurée) pour régler la valeur dialnorm de l'étape suivante d'encodage AC-3. À cette étape de recodage, il est crucial que les autres paramètres de métadonnées soient réglés uniformément et de façon appropriée.

9 GESTION DE LA DYNAMIQUE

Le système audio de TVN est capable de fournir une gamme dynamique (intervalle entre les sons les plus faibles et les sons les plus forts) très large. Les producteurs de contenu se servent souvent de la gamme dynamique comme une des méthodes de communication de l'intention artistique.

Toutefois, il pourrait y avoir un conflit entre le producteur désirant distribuer un contenu de gamme dynamique étendue et le téléspectateur qui ne peut pas, ou ne veut pas profiter de cette gamme dynamique accrue. L'impossibilité peut provenir de l'équipement du téléspectateur qui n'est pas en mesure de reproduire la gamme de sons voulue ou d'un environnement d'écoute qui ne permet pas de jouir de la gamme dynamique étendue. Ainsi, l'objectif de conserver la gamme dynamique originale du contenu et celui de satisfaire les téléspectateurs ne vont souvent pas de pair.

Un des objectifs du système AC-3 est de doter les producteurs de contenu de la plus grande liberté et de la plus grande souplesse dans le choix du contrôle de la gamme dynamique (DRC) au moment de créer leur contenu. Le système AC-3 transmet ces options de DRC au téléspectateur, dont les choix agiront avec le système DRC d'une façon connue et reproductible.

Il existe plusieurs méthodes de contrôle de la gamme dynamique. Certaines sont appliquées avant l'encodage du signal audio, d'autres, après le décodage et d'autres encore, dans les deux domaines :

- 1) Une approche est la compression classique et (ou) la limitation, où le réglage du gain est appliqué aux signaux audio avant l'encodage.

- 2) Une autre approche est l'emploi du système de codage AC-3, qui génère des mots de contrôle de gain durant l'encodage, mais qui n'applique le réglage du gain aux signaux audio qu'après le décodage, ce qui permet à l'utilisateur de choisir facultativement l'étendue de la plage dynamique qu'il souhaite.

La principale différence entre les deux approches est que la méthode AC-3 produit un résultat « réversible » et l'autre, un résultat permanent. Une méthode hybride peut aussi être adoptée, qui combine une partie de traitement permanent et une partie « réversible » dans une proportion établie par le radiodiffuseur.

9.1 Système de contrôle de dynamique AC-3 (« réversible »)

Pour donner la possibilité à la plus grande partie des téléspectateurs d'apprécier le contenu dans la plus grande diversité de conditions d'écoute, un mécanisme intégré au système audio de TVN est conçu pour permettre de restreindre la gamme dynamique du contenu. En termes simples, le système de contrôle de dynamique (DRC) AC-3 peut être considéré comme un compresseur/limiteur divisé en deux; le signal de contrôle est généré dans l'encodeur et le signal de contrôle peut être appliqué dans le décodeur. Dans ce système, le signal audio n'est pas modifié de l'original avant d'atteindre le décodeur du consommateur, de sorte que le contrôle de la gamme dynamique peut être considéré comme « réversible ». Les détails du fonctionnement du DRC sont présentés à l'annexe F.

Les métadonnées sont toujours transmises comme partie du flux audio codé. Les encodeurs AC-3 nécessitent le réglage précis de certains paramètres de métadonnées. Le DRC est un de ces paramètres et *dialnorm*, ou le paramètre d'intensité sonore indiquée, en est un autre. Bien qu'il soit distinct de *dialnorm*, le bon fonctionnement du système DRC est très dépendant d'une mesure juste du contenu et de l'indication également juste de l'intensité sonore de ce contenu. On trouvera les lignes directrices sur les mesures dans la section 5 du présent document.

Le système DRC modifie la gamme dynamique du contenu décodé en abaissant le niveau des parties très « fortes » du contenu pour éviter de gêner le téléspectateur et en élevant le niveau des parties très « faibles » pour en améliorer l'adaptation à l'environnement d'écoute. Si, toutefois, le téléspectateur a un cinéma maison ou autre environnement d'écoute pouvant reproduire toute la gamme dynamique, il peut choisir de mettre hors fonction le système DRC pour recevoir le contenu dans sa version mixée originale.

Quand le téléspectateur met hors fonction le système DRC, la gamme dynamique du contenu est identique à la dynamique d'entrée dans l'encodeur AC-3. Étant donné que l'utilisation du système DRC est sélectionnable par le téléspectateur, le système AC-3 peut fournir plusieurs options d'écoute au téléspectateur, pour répondre à divers systèmes de reproduction et environnements d'écoute.

9.1.1 DRC en mode ligne et en mode RF

Dans l'encodeur audio AC-3, deux signaux de commande sont calculés par le système DRC :

- 1) Un signal de commande est employé pour réduire légèrement la gamme dynamique du contenu (convient aux composantes audio directement raccordées).
- 2) L'autre signal de commande est employé pour le contrôle plus énergique de la gamme dynamique (convient au matériel raccordé avec signaux analogiques modulés RF ou au matériel autrement incapable de gérer la gamme dynamique étendue potentielle du signal original).

Le fabricant du matériel choisit l'un de ces deux signaux comme mode par défaut, en fonction de la conception et de l'utilisation prévue du produit. Pour les téléspectateurs désirant la gamme dynamique intégrale du signal original, il est possible de ne pas tenir compte de ces signaux de commande et d'obtenir la reproduction du contenu exactement comme le producteur l'a créé. Cette méthode d'emploi de signaux de commande et audio distincts résulte en un système DRC « réversible », capable de satisfaire simultanément des téléspectateurs différents, dans des conditions d'écoute et des situations différentes.

Deux types de signaux de commande DRC sont calculés, car il existe deux scénarios suivant lesquels la réduction/l'augmentation de gain peut être souhaitée.

Dans la première situation, une réduction légère de la gamme dynamique est voulue. C'est le mode par défaut de nombreux dispositifs qui utilisent les interconnexions analogiques « bande de base » (souvent disponibles sur les sorties rouge et blanche des connecteurs de sortie « RCA »). Ces connexions ne constituent pas une interface professionnelle, mais sont souvent acceptables comme interface consommateur. Ce signal de commande DRC est appelé « mode ligne » dans l'interface de commande d'encodeur dans la plupart des produits professionnels, car il fait souvent référence au processus d'interconnexion au niveau lignes. Le nom du paramètre de métadonnées correspondant est `dynrng`.

Un deuxième signal de commande DRC, plus énergique, est calculé en même temps. Il est conçu pour l'emploi dans les produits pouvant être interconnectés par un signal analogique modulé RF. La gamme dynamique des signaux modulés RF est généralement plus étroite que celle des signaux utilisant le processus d'interconnexion au niveau lignes. En mode RF, un gain de 11 dB est appliqué aux fins d'harmonisation de l'intensité sonore avec celle des anciens systèmes analogiques. Ce mode est le mode par défaut des décodeurs de MVPD ainsi que des téléviseurs équipés de syntoniseurs numériques et de petits haut-parleurs intégrés. Le nom du paramètre de métadonnées correspondant est `compr`. Pour obtenir plus de renseignements à propos des recommandations visant l'harmonisation de l'intensité sonore pour les produits grand public, voir le document CEA-CEB11 [25]. Les recommandations du présent document et du document CEA-CEB11 ont été coordonnées et sont complémentaires.

9.1.2 Contrôle

L'emploi d'un système dans lequel les « mots de gain » (voir annexe F) de DRC sont calculés dans l'encodeur et appliqués dans le décodeur présente des avantages. Un des avantages de ce type de conception système est qu'il permet de visualiser, ou d'émuler, les fonctions de compresseur/limiteur en cours de production, bien avant que le contenu soit codé. Les producteurs peuvent vérifier le contenu au fur et à mesure qu'il est produit pour en entendre le résultat selon des modes de dynamique préétablis.

Les effets de ce système DRC seront audibles aux mixeurs travaillant dans un environnement de contrôle professionnel, mais le processus devrait être considéré dans le contexte du consommateur type, dans un environnement domestique type, où la réduction et l'augmentation de gain dont il est question passent généralement inaperçues. Dans la plupart des situations, les effets du DRC en mode ligne améliorent le rendu général du contenu en adaptant mieux les signaux audio aux capacités de reproduction de l'équipement et de l'espace d'écoute du téléspectateur.

9.1.3 Relations avec `dialnorm`

Tous les calculs de DRC sont fondés sur l'intensité sonore indiquée du contenu et relatifs à cette intensité, représentée par le paramètre de métadonnées `dialnorm`. En d'autres mots, l'encodeur doit

connaître le niveau sonore prévu du contenu pour déterminer quand le contenu est « trop fort » ou « trop faible ». Le paramètre `dialnorm` définit efficacement ce niveau cible. En conséquence, il est très important que `dialnorm` indique avec précision l'intensité sonore du contenu.

Le paramètre `dialnorm` est aussi utilisé pour établir le seuil d'un processus de protection contre les surcharges quelque peu dissimulé et inévitable, conçu pour empêcher les signaux audio réduits par mixage de surcharger le matériel du consommateur. La protection contre la surcharge utilise une balistique appropriée pour éliminer les risques de surcharge, mais elle est beaucoup moins qu'idéale pour la qualité sonore. Il convient d'éviter la protection contre les surcharges. Il est possible de le faire en veillant à ce que le paramètre `dialnorm` représente avec précision l'intensité sonore réelle du contenu audio.

9.1.4 Codage professionnel

Dans les encodeurs AC-3, les caractéristiques de réduction et d'augmentation de gain des modes DRC RF et de ligne sont déterminées par un groupe de « profils » DRC. Ces profils décrivent de nombreux paramètres, y compris la plage de réduction de gain, la plage d'augmentation de gain ainsi que les temps d'attaque et de relâchement. Entre ces plages se trouve une plage linéaire (dite « zone sans correction ») où aucune réduction ou augmentation de gain n'a lieu. Il est prévu que la majorité des contenus mixés de manière professionnelle se trouvera dans la plage « sans correction », où le contenu sera distribué exactement tel qu'il est produit, sans modification supplémentaire (ou continue). Les excursions au-delà de cette zone sans correction peuvent servir à communiquer une intention artistique particulière. Il est à noter que le paramètre `dialnorm` détermine la position de la zone sans correction, ce qui indique, une fois de plus, qu'il est essentiel que le paramètre `dialnorm` représente avec précision l'intensité sonore du contenu.

Cinq profils sont définis dans l'encodeur AC-3, qui sont les suivants :

- Music Light (Musique léger)
- Music Standard (Musique standard)
- Film Light (Film léger)
- Film standard
- Speech (Parole).

La différence entre les options de DRC musique et film peut être subtile pour un auditeur type, mais l'un ou l'autre peut convenir davantage à certains types de contenu. La meilleure façon de le déterminer est de contrôler au moyen d'un émulateur approprié. Les versions « légères » des profils ont une zone sans correction beaucoup plus large. Ainsi, la réduction ou l'augmentation de gain commence plus loin du niveau audio moyen du programme, ce qui réduit la réduction ou l'augmentation de gain par rapport à la version « standard » du profil.

Comme son nom l'indique, le profil parole est prévu pour les programmes ne contenant que de la parole (un format de station « radio interactive », par exemple). Ce profil peut introduire un phénomène parasite de DRC perceptible dans les programmes ayant de la musique et des effets sonores. Pour obtenir plus de renseignements sur les profils DRC, veuillez consulter l'Annexe F.

9.1.5 DRC AC-3 : sélection de « None »

Il y a aussi un choix appelé « None » (aucun), qui a pour effet de ne sélectionner aucun des profils DRC nommés. La sélection de l'option « None » (par l'opérateur) empêche la création des mots de commande DRC.

L'option « None » est un choix acceptable dans la mesure où l'on comprend bien les implications de ne pas choisir de profil DRC :

- La caractéristique de réversibilité du système DRC ne sera pas offerte au consommateur.
- La sélection de « None » empêche le téléspectateur de choisir une option DRC ou d'activer des fonctions comme « Late Night » (fin de soirée) ou « Midnight » (nuit) dans certains appareils qui utilisent le mode DRC RF.
- La gamme dynamique devrait être contrôlée d'une autre façon par l'opérateur ou par le créateur du programme.
- Dans certains récepteurs de TVN ayant une capacité de volume limitée, il est possible que la limite de reproduction soit dépassée.
- Les mots de commande DRC du mode RF sont aussi utilisés pour établir des limites de protection. Un limiteur de protection empêche l'écrêtage dans les décodeurs des consommateurs, qui pourrait être entraîné par un réglage $dialnorm$ erroné conjugué avec une programmation très dynamique. La sélection de l'option « None » n'empêche pas la création de mots de commande DRC de limite de protection, qui sont très énergiques et peu artistiques. Le processus de limitation de protection a un temps d'attaque très court et un temps de relâchement très long, ce qui peut produire des phénomènes parasites audibles désagréables.
- Les systèmes qui utilisent les signaux ATSC comme source de distribution DS (p. ex. câblodistribution analogique DS) ne seront pas en mesure d'employer le DRC en mode RF pour établir un signal analogique DS de gamme dynamique réduite. Pour plus de conseils, voir la pratique recommandée ATSC A/79 [24].

Pour que le système de contrôle de gamme dynamique AC-3 soit fonctionnel, les opérateurs devraient choisir un autre profil que « None » lorsqu'approprié.

9.2 Traitement de la dynamique avec interface de métadonnées (« réversible »)

Des processeurs de métadonnées audio ont été développés, qui permettent la création de plusieurs valeurs de gain en mode ligne et en mode RF pouvant être enregistrées dans des pré-réglages ou profils définis par l'utilisateur. Il est possible d'exécuter un traitement supplémentaire dans le domaine des métadonnées pour obtenir des résultats beaucoup plus proches de ceux du traitement classique, mais avec l'avantage de la « réversibilité ». Ces technologies fonctionnent de pair avec l'encodeur AC-3, mesurant les signaux audio et calculant les valeurs de gain, puis les passant à l'encodeur pour insertion dans le flux binaire AC-3. L'application est analogue aux fonctions de ligne et RF décrites ci-dessus, mais elle a recours à des profils différents pour remplir des objectifs de traitement de la dynamique différents.

Ces profils de remplacement permettent de prévoir un rapport traitement DRC permanent-traitement DRC par métadonnées réglable. À un extrémité de l'ajustement, le traitement audio multibande permanent est appliqué au flux audio après qu'il a été d'abord préconditionné, par application des métadonnées, pour réduire au minimum le degré de traitement permanent pouvant être nécessaire.

Les signaux audio résultants sont encodés AC-3 avec une valeur $dialnorm$ statique et l'un des profils DRC classiques peut être sélectionné. À l'autre extrémité du rapport réglable, le flux audio qui a été préconditionné par application de métadonnées est ensuite analysé et les mots de commande de gain DRC pour le mode de ligne et le mode RF sont générés par sélection par l'opérateur.

Les paramètres classiques tels que temps d'attaque (*attack*) et de relâchement (*release*), seuil (*threshold*), plage de commande automatique de gain (CAG), fenêtre antiparasite (*gate*) et

maintien (*freeze*) peuvent être réglés de manière très semblable à celle d'un processeur audio permanent, mais les résultats sont plutôt inclus dans le flux binaire AC-3, accompagnant le programme audio codé. Il est ainsi possible de réaliser un traitement plus énergique que celui qui est prévu dans les cinq profils d'origine. Le programme audio original est livré au consommateur et les valeurs DRC sont appliquées au signal audio par défaut, mais le téléspectateur peut les désactiver s'il le souhaite.

9.3 Traitement de la dynamique sans interface de métadonnées (irréversible)

En diffusion radio AM analogique, les systèmes de CAG ont été créés pour assurer la modulation appropriée de la fréquence de porteuse des stations de radiodiffusion. Une modulation erronée pouvait réduire la couverture si le niveau moyen était trop bas et pouvait provoquer de l'interférence si le niveau de crête était trop élevé. Cette opération est devenue plus difficile pour les systèmes FM, qui utilisent la préaccentuation pour augmenter le niveau des hautes fréquences avant l'émission afin de réduire le bruit au minimum. Essentiellement, ces systèmes de CAG (souvent appelés « contrôleurs de modulation ») contrôlaient le rapport valeur de crête sur valeur moyenne des signaux audio, et donc, la dynamique. À la différence des approches à métadonnées décrites plus haut, les modifications de gain sont appliquées immédiatement au signal d'entrée, avant la transmission aux consommateurs, et les résultats sont donc jugés irréversibles.

Les premiers systèmes de CAG étaient des systèmes large bande comportant un seul détecteur de niveau et un seul élément de gain pour contrôler toute la largeur de bande audio. Dans les débuts de la radiodiffusion, cette pratique pouvait être acceptable, parce que la largeur de bande audio était limitée et que des opérateurs compétents réalisaient des programmes audio plus uniformes. Toutefois, en raison de l'augmentation de la largeur de bande audio et de la diminution de l'intervention des opérateurs, il fallait développer d'autres techniques pour maintenir le contrôle et réduire au minimum les effets secondaires audibles du processus de CAG.

Un phénomène parasite potentiel de la CAG large bande est la possibilité que le niveau d'une partie du spectre audio soit modifié à cause d'une autre partie du spectre audio sans rapport avec la première, qui a plus de puissance. Pour représenter ce phénomène, on peut imaginer un signal audio contenant les sons d'une grosse caisse basse fréquence et d'une flûte de fréquence plus élevée. En réponse à l'énergie élevée de la grosse caisse, le système de CAG large bande simple ferait varier le signal de commande, et donc le niveau, de toutes les fréquences. Cette modification de niveau pourrait être appropriée pour la grosse caisse, mais entraînerait une fluctuation inappropriée dans le niveau du signal de la flûte, un effet communément appelé « pompage ». Le développement de systèmes de traitement multibandes a remédié largement à ce problème, car le signal d'entrée est divisé en deux bandes de fréquences ou plus et une CAG est appliquée à chacune des bandes. Les systèmes multibandes permettent un plus grand contrôle accompagné de moins d'interaction entre les différentes parties du spectre audio, mais peuvent aussi modifier l'équilibre spectral du signal appliqué.

D'autres techniques ont évolué pour réduire au minimum les effets secondaires comme le « pompage », la surcharge entraînée par la préaccentuation et l'augmentation du bruit de fond. Il s'agit notamment du traitement en série, du traitement en prélecture, de l'écrêtage et du fenêtrage antiparasites. Le traitement en série consiste à mettre en série deux sections de CAG ou plus, de vitesse différente, pour contrôler séparément les crêtes et les signaux de niveau moyen. Le traitement en prélecture prévoit un délai entre les parties détection et réglage de la CAG de manière à ce que les modifications de gain soient appliquées au moment précis où il le faut.

L'écèlement est un processus au cours duquel les crêtes des signaux sont tronquées; il peut entraîner une distorsion audible. Les fenêtres antiparasites empêchent les signaux de très faible niveau et le bruit d'être inutilement amplifiés par la CAG et peuvent être utilisées avec des seuils d'autres phases pour créer une zone sans traitement. Ensemble, toutes ces techniques dépassent considérablement les objectifs initiaux de protection simple contre la sous-modulation ou la surmodulation et dans de nombreux cas, elles ont été utilisées pour produire un son très différent de l'original.

Le but des systèmes modernes de contrôle de gamme dynamique en télévision numérique est différent de celui de leurs prédécesseurs analogiques. L'élimination de la nécessité de préaccentuation réduit énormément la quantité de traitement nécessaire. Les appareils récents peuvent aussi faire appel à des modèles de détection complexes, fondés sur la recommandation UIT-R BS.1770 [3] ou compatibles avec elle. Si l'objectif est simplement de gérer l'intensité sonore en produisant une incidence minimale sur le contenu des programmes, l'expérience a démontré qu'un processus à plusieurs phases de traitement subtil peut être très efficace tout en permettant de conserver dans une large mesure l'intégrité des programmes. En outre, il est possible de régler les problèmes survenant aux extrémités des programmes et du matériel interstitiel au moyen d'un traitement de gamme dynamique irréversible, qui ne fait pas partie du système AC-3⁶.

9.4 Expérience du consommateur

Les téléspectateurs peuvent agir de plusieurs façons sur le processus de DRC à l'intérieur des flux binaires AC-3. En règle générale, les décodeurs des MVPD et les téléviseurs numériques à syntoniseur/décodeur intégré fonctionnent dans un mode DRC par défaut. Il s'agit généralement du mode RF, car ces appareils contiennent un modulateur RF ou alimentent directement des téléviseurs qui ne permettent qu'une gamme dynamique minimale. Souvent, le prix et la souplesse de l'équipement déterminent le nombre d'options dont dispose le téléspectateur.

Les récepteurs audio-vidéo de cinéma maison présentent souvent le plus grand nombre d'options d'écoute DRC, mais les menus intégrés pour naviguer dans la sélection peuvent manquer de cohérence et d'interopérabilité avec d'autres produits. Dans ce contexte, manque d'interopérabilité signifie que la terminologie employée pour décrire les options peut différer d'un appareil à un autre, voire être contradictoire. Par exemple, les notions de « wide » (étendue), « standard » (ou « normal ») et « narrow » (étroite) sont couramment utilisées. Ici, « wide » signifie qu'aucun DRC n'est activé et que le résultat est la reproduction de la gamme dynamique étendue du produit original. « Standard » signifie l'application du profil sélectionné pour le mode ligne durant l'encodage. « Narrow » signifie l'application du profil sélectionné pour le mode RF durant l'encodage. Comme nous l'avons mentionné plus tôt, en mode RF, une accentuation du signal de 11 dB est appliquée pour faire correspondre l'intensité sonore à celle des anciens systèmes analogiques.

10 CONFIGURATION DE CONTRÔLE AUDIO

La présente section décrit la configuration recommandée des systèmes audio pour la télévision numérique, y compris l'alignement des systèmes de contrôle de la régie sur un niveau de pression acoustique de référence, correspondant à un niveau électrique particulier, ainsi que l'égalisation de ces systèmes audio. Les conditions de contrôle visant plusieurs espaces utilisés pour la production et la postproduction télévisuelles sont indiquées dans le tableau 10.1.

⁶ Pour obtenir plus de renseignements, consulter les fabricants d'appareils conçus à cette fin.

Cette section n'est pas destinée à remplacer la recommandation SMPTE RP 200 [6] pour le contrôle sonore cinématographique dans de plus grands espaces.

Elle décrit les conditions de configuration initiale recommandées pour le contrôle, ainsi que les mesures et les réglages devant être exécutés de temps à autre aux fins de contrôle de qualité du contrôle dans les espaces utilisés pour la production télévisuelle et le contrôle de la qualité. Elle s'adresse aux planificateurs, aux ingénieurs concepteurs, aux techniciens d'installation, aux techniciens de maintenance et aux utilisateurs.

Un « aide-mémoire » pour le réglage des niveaux de contrôle est présenté à l'annexe D.

Note : Les signaux d'essai indiqués dans le présent document sont tirés de la série de disques *TMH Test Disc Series* distribuée par The Hollywood Edge et sont protégés par droit d'auteur. L'utilisation gratuite est accordée aux studios de vidéo et de télévision, aux réseaux et aux maisons de postproduction aux fins d'alignement. Tous autres droits réservés. La série *TMH Test Disc Series* propose de nombreux autres signaux d'essai pour les essais acoustiques et électriques.

10.1 Contexte

Comme nous l'avons signalé dans la section 1.1, les consommateurs ne s'attendent pas à percevoir de grandes variations d'intensité sonore au passage d'un programme à un contenu interstitiel et quand ils changent de chaîne. Deux fonctions liées au niveau sonore sont intégrées au système AC-3 spécifié dans la norme ATSC : la normalisation du dialogue et le contrôle de la gamme dynamique. Si ces fonctions sont bien réglées et bien utilisées, elles permettent d'atteindre dans une large mesure les objectifs de maintien de l'intégrité artistique des mixages, tout en livrant une gamme dynamique convenant aux conditions d'écoute du consommateur. L'utilisation d'une disposition de haut-parleurs de contrôle normalisée, de niveaux de référence électriques et acoustiques standard et d'une réponse spectrale de référence par tous les fournisseurs de programmes et de contenu interstitiel produit une base commune pour le contrôle. Outre les fonctions intégrées au système AC-3, l'adoption pratique des recommandations de cette section de la PR devrait contribuer à satisfaire aux besoins réciproques des producteurs et des consommateurs.

L'observation universelle est qu'un niveau de pression acoustique donné est perçu comme étant plus fort dans de petits locaux, comme une régie, que dans de grandes salles comme les cinémas. L'interchangeabilité du niveau de pression acoustique de référence utilisé dans ce document avec le niveau SMPTE RP 200 [6] employé dans les grands espaces a été testée.

Il a été démontré que l'utilisation de conditions de contrôle de référence améliore l'interchangeabilité des programmes, des messages commerciaux et d'autres mixages interstitiels et qu'elle est donc souhaitable.

10.2 Caractéristiques des locaux et des espaces

Cinq catégories de régies audio et d'espaces de postproduction sont définies. Elles sont présentées dans le tableau 10.1, accompagnées de leurs caractéristiques.

Tableau 10.1 Catégories de régies audio utilisées en production télévisuelle

Catégorie	Caractéristiques
I	Régie audio principale, dotée d'une acoustique et de systèmes audio spécialisés. Gamme de canaux pouvant atteindre 5.1 (3 avant/2 ambiance/0,1 enrichissement basses fréquences). Bien isolée d'autres

	activités. Gamme de fréquences et dynamique les plus étendues, égales aux meilleurs cinémas maison alignés de façon appropriée. Ce type de local peut être utilisé pour le contrôle de la qualité à l'échelle du réseau, par exemple pour vérifier la conformité du matériel aux exigences de distribution quand une question est soulevée dans les stations d'ingestion. La qualité du contrôle audio dépasse les exigences de production dans les locaux de cette catégorie. On peut s'attendre à ce que les entreprises de radiodiffusion n'aient qu'un petit nombre de ces locaux.
II	Espace de production avant tout audio, ayant des besoins d'équipement et un positionnement qui remplacent les conditions absolues de contrôle audio, bien que le contrôle audio y soit en principe bon. Nombre de canaux égal au plus grand nombre utilisé pour le matériel créé dans le local. Bonne isolation des autres activités. Ce type de local peut être utilisé pour la création de programmes, les produits qui en sortent devant à l'occasion être soumis à un contrôle dans un local de catégorie I. La gamme de basses fréquences et la marge de sécurité peuvent être quelque peu limitées par rapport à celles permises dans un local de catégorie I.
III	Espace de montage audio, locaux de prémixage parole et de prémixage effets-ambiance et autres espaces dont il est généralement prévu que les produits soient intégrés aux programmes créés dans un local de catégorie II ou mieux. Si l'espace est utilisé pour le mixage final, appliquer la pratique recommandée pour les niveaux et l'égalisation indiquée dans le présent document.
IV	Camions et cabines de mixage de programmes. Ces espaces exigent des mesures particulières en raison de leur petit volume, du niveau élevé de bruit de fond, d'un niveau élevé de premières réflexions et des besoins de communication dans un environnement de production.
V	Recommandations visant les systèmes de contrôle à casque d'écoute. Utilisés dans les stations « d'ingestion », dans des environnements encombrés, pour le contrôle de qualité dans les salles de machinerie et dans les endroits semblables.

10.3 Installation

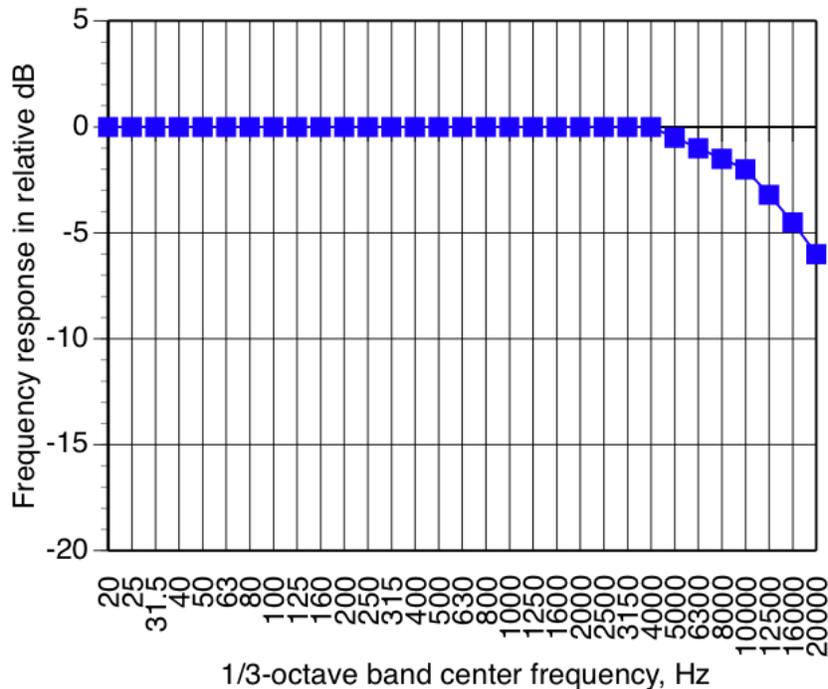
Les systèmes audio des espaces de catégories I, II et IV devraient être installés, pour autant que ce soit réalisable, suivant les pratiques recommandées dans les documents de référence (voir références [12], [14], [16] et [17]). Les espaces de catégorie III devraient aussi remplir les exigences des pratiques recommandées s'ils sont utilisés pour le mixage final.

Il convient de souligner l'importance de placer les haut-parleurs à égale distance de l'emplacement de contrôle de mixage principal ou, si ce n'est pas possible, d'utiliser une temporisation de façon à ce que le temps d'arrivée à l'emplacement de mixage soit constant pour tous les canaux de haut-parleur. Cette mesure est nécessaire en raison de la très grande précision de l'ouïe humaine dans la localisation d'images fantômes apparaissant entre les paires adjacentes de haut-parleurs. Cette exigence est particulièrement importante pour les canaux avant et, si les haut-parleurs d'ambiance sont plus près de l'emplacement d'écoute que les haut-parleurs avant, il convient d'y appliquer une temporisation.

Les documents de référence présentent la réponse en fréquence anéchoïque prévue des haut-parleurs et des systèmes de contrôle. Des mesures des haut-parleurs sur place, dans les régies, indiquent toutefois de forts écarts dans la réponse anéchoïque des haut-parleurs, particulièrement en raison des conditions de charge aux zones limites des salles à très basse fréquence, accompagnés d'effets modaux d'onde stationnaire, généralement dans la gamme de fréquences allant de 80 Hz à 500 Hz. C'est pourquoi l'égalisation de la salle est hautement souhaitable et même nécessaire pour les espaces de qualité supérieure. Les problèmes liés à l'égalisation des salles sont décrits à l'annexe C.

Il a été observé que les plus grands écarts en balance octave à octave des mixages diffusés se situent dans les fréquences des extrémités, sous 100 Hz et au-dessus de 8 kHz. Ce phénomène est probablement imputable aux réponses variables des systèmes de contrôle dans ces gammes de fréquences. C'est pourquoi la courbe de réponse dans une salle fonctionnelle présentée à la figure 10.1 s'applique aux espaces de contrôle de la catégorie I. La gamme basses fréquences et

la marge de sécurité des espaces de catégories II et IV peuvent être réduites par rapport à celles des espaces de catégorie I.



interchangeables dans diverses conditions d'écoute. Les étapes devraient être exécutées comme suit.

Étape 1. Exécuter l'alignement électrique du système à étalonner au moyen de l'essai suivant :

- Signal 1. Tonalité sinusoïdale 440 Hz à -20 dB FS enregistrée sur le canal gauche de ce fichier stéréo. Importer le fichier dans le poste de travail audio numérique et le dupliquer pour chaque canal.
http://www.atsc.org/refs/a85/440Hz_left_Ch-20dB.wav.
- Utiliser cette tonalité pour aligner les indicateurs de sortie de l'équipement utilisé pour -20 dB FS sur les appareils de mesure numériques. Utiliser un appareil de mesure de résolution appropriée. La meilleure configuration est que l'appareil de lecture et les indicateurs d'entrée, de canal et de sortie du pupitre puissent tous être réglés avec précision sur le gain unitaire.
- Sur les appareils de mesure analogiques, ce niveau est réglé sur 0 VU (ANSI C16.5-1942 [25], IEC 60268-17 [27]), sur 4 à l'échelle de 7 du crête-mètre BBC (IEC 60268-10/Ia, IIb: 1991 [28]) (bien qu'il soit à noter que ce crête-mètre affiche 4 dB/pas, de sorte que la valeur étalonnée maximale correspond à -8 dB FS; ce problème est atténué par le fait que le crête-mètre PPM a un temps d'attaque de 80 pour cent de l'indication en 10 ms, beaucoup plus lent qu'un vrai crête-mètre) et sur les valeurs correspondantes pour les autres variantes d'appareils de mesure.
- Pour plus de détails sur ce signal d'essai, voir la section 10.5.1.

Étape 2. Si on ne dispose pas d'un sonomètre, passer à l'étape 3. Aligner le niveau sonore du système à étalonner au moyen de l'essai suivant :

- Signal 2. Bruit rose limité à la bande 500 Hz - 2 kHz à -20 dB FS enregistré sur le canal gauche de ce fichier stéréo. Importer le fichier dans le poste de travail audio numérique et le copier sur les canaux gauche, droite, centre, ambiance gauche et ambiance droite, un à la fois et tour à tour. Maintenir la structure de gain unitaire établie avec la tonalité sinusoïdale. Pour plus de détails sur ce signal d'essai, voir la section 10.5.3.
http://www.atsc.org/refs/a85/MidRngPinkNoise_-20dB.wav
- Ne pas tenir compte des indications du pupitre ou d'autres appareils de mesure électriques, qui seront différentes pour le bruit et pour la tonalité, car le niveau de crête de cette source de bruit stochastique (aléatoire) est plus élevé que sa valeur quadratique moyenne, qui est elle-même plus élevée que le niveau moyen. Sur les crêtes-mètres, la plage se situe autour de 10 dB de plus que la tonalité sinusoïdale, alors que sur le vumètre (qui est un instrument de réponse moyenne étalonné sur la valeur efficace) conforme à IEEE C16.5-1954, le niveau indiqué sera environ 1 dB plus bas.
- Appliquer ce signal sur chaque canal, tour à tour, et ajuster les commandes appropriées ne touchant que le niveau de contrôle, sur le niveau acoustique standard. Dans n'importe quelle installation, les éléments pouvant toucher le niveau acoustique de contrôle sont la commande de niveau de sortie de contrôle du pupitre, qu'il est suggéré de régler sur un standard marqué, ainsi que les commandes de niveau d'égaliseur de salle/haut-parleurs et les commandes de gain d'amplificateur de puissance ou commandes de sensibilité d'entrée de haut-parleurs amplifiés pouvant être utilisés. Les problèmes ordinaires de marge de sécurité et de rapport signal-bruit dans une chaîne de plusieurs appareils s'appliquent.

Le niveau de pression acoustique standard devrait être mesuré au moyen d'un sonomètre, de préférence conforme au type 2 de la norme ANSI S1.4-1983 [29], en mode de mesure lente (temps d'intégration de 1 s) et avec pondération C. Il devrait être mesuré en un point correspondant au centre de la tête d'un auditeur assis à l'emplacement de mixage principal, le sonomètre étant orienté vers le canal à étalonner. Il est à noter que le corps de la personne effectuant la mesure devrait se trouver d'un côté de l'appareil de mesure, non derrière l'appareil par rapport à la source, où la réflexion des ondes peut modifier la mesure.

Les niveaux de référence acoustique normalisés pour chaque catégorie et pour des locaux de divers volumes à l'intérieur des catégories sont présentés dans le tableau 10.2.

Tableau 10.2 Niveau de pression acoustique de référence

Catégories	Volume du local en pieds cubes	Niveau de pression acoustique en dB relatifs à 20 $\mu\text{N}/\text{m}^2$
I, II	> 20 000	85*
	10 000 < 19 999	82
	5 000 < 9 999	80
	1 500 < 4 999	78
	< 1 499	76
III	Est fonction de l'utilisation du local. Pour le montage, peut être ajusté par le monteur selon l'utilisation du matériel dont il dispose. Pour le mixage final de programme, suivre les recommandations visant les catégories I et II ci-dessus.	
IV	< 1 500	76
V		Utiliser un coupleur 2 cm ³ et régler le niveau du signal 440 Hz à 74 dB.
* Conformément à SMPTE RP 200 [6]		

Étape 3. Repiquer le fichier de parole original téléchargé (voir ci-dessous), en gain unitaire, dans le canal du centre ou mono de l'appareil de lecture utilisé. Dans le cas où l'on ne peut exécuter que le contrôle stéréo sur deux canaux, reproduire le fichier de parole original dans les canaux gauche et droit avec atténuation de -3 dB dans chaque canal et confirmer qu'ils sont synchronisés dans l'échantillon. Veiller à ce que le trajet du signal soit le même qu'à l'étape 1 ci-dessus, de manière à ce que cette lecture se fasse dans les mêmes conditions de gain que l'étape 1 et de bruit que l'étape 2. Faire jouer l'enregistrement de parole, dont l'intensité sonore mesurée est de -24 LKFS et observer si le niveau est normal. Si on ne dispose pas de sonomètre et que l'on a passé l'étape 2, s'assurer d'avoir exécuté l'étape 1 si l'on dispose de l'équipement indiqué et ajuster cette piste pour obtenir le niveau d'écoute le plus confortable possible.

- Le niveau sonore de cet enregistrement de parole a été réglé à -24 LKFS, conformément à la norme BS.1770 [3].

http://www.atsc.org/refs/a85/Speech_sample.wav

Étape 4. S'il y a un caisson d'extrêmes graves, alimenter le canal d'effets basses fréquences électriquement au moyen du signal ci-dessous et régler le niveau du caisson pour $+4$ dB par rapport aux canaux principaux, quand on exécute la mesure au moyen d'un sonomètre à lecture lente et pondération C.

http://www.atsc.org/refs/a85/LFPinkNoise_-20dB.wav

10.5 Détails des signaux d'essai

10.5.1 Définition de 0 dB FS

L'incertitude des mesures de niveau de bruit au moyen d'appareils de mesure d'électricité est décrite ci-dessus à la section 10.1. La définition de niveau dans les systèmes audio numériques est indiquée dans le document AES17-1998 (R2004) [21], où une onde sinusoïdale pleine échelle est définie à 0 dB FS; cette définition s'applique dans le présent document. Toutefois, de nombreux logiciels indiquent le niveau sur des appareils virtuels en se fondant sur des calculs de valeurs efficaces classiques, donnant une indication d'onde sinusoïdale pleine échelle de $-3,01$ dB FS, ce qui est erroné dans le contexte du présent document.

10.5.2 Tonalité sinusoïdale 440 Hz

http://www.atsc.org/refs/a85/440Hz_left_Ch-20dB.wav

La tonalité sinusoïdale est le signal d'essai audio le plus simple et le plus largement employé. Les utilisations types comprennent le contrôle du cheminement des signaux et le réglage du niveau. Une onde sinusoïdale présente deux caractéristiques utiles pour les essais : premièrement, elle est de niveau constant, ce qui permet de la mesurer avec une plus grande précision et deuxièmement, une tonalité sinusoïdale de fréquence moyenne n'impose aucune contrainte aux limites des systèmes et les traverse sans subir de traitement.

Le niveau est de -20 dB par rapport au maximum (dB FS).

La fréquence de 440 Hz a été choisie pour trois raisons : elle se situe dans la portion la plus plate de la courbe de réponse en fréquences d'égalisation de la norme BS.1770 [3], elle est musicalement pertinente comme référence « A » sur l'échelle musicale et elle n'est pas harmoniquement liée aux diverses fréquences d'échantillonnage utilisées, ce qui permet l'exploitation de toutes les valeurs de code de sa portée. Utiliser une tonalité à une fréquence de syntonie précise facilite les essais en permettant d'identifier de façon audible les erreurs pouvant exister dans la fréquence d'échantillonnage de lecture par rapport à la fréquence d'échantillonnage enregistrée.

Durant le développement de ce signal d'essai, on a découvert que la production d'une onde sinusoïdale d'essai à précisément -20 dB FS, suivie de l'ajout d'un bruit blanc vibratoire de fonction de densité de probabilité triangulaire de ± 1 LSB, comme il est nécessaire pour obtenir un signal d'essai sans distorsion, déclenchait dans certains appareils de mesure une indication d'une graduation au-dessus de -20 dB FS, pouvant produire une erreur atteignant 2 dB dans ces appareils. Le niveau a donc été réduit d'un bit en amplitude, puis la vibration ajoutée, de sorte que le niveau de crête du signal est précisément -20 dB FS, bruit inclus.

10.5.3 Bruit rose limité à la bande

http://www.atsc.org/refs/a85/MidRngPinkNoise_-20dB.wav

Un bruit rose à bande limitée est préféré à un bruit à bande intégrale pour plusieurs raisons. À basse fréquence, sous la fréquence de Schroeder⁷ dans les salles, les ondes stationnaires ont une forte influence sur le niveau et l'inclusion de telles fréquences dans la bande de mesure ajoute une incertitude. De plus, le contenu énergétique des basses fréquences dans un signal stochastique entraîne de plus grandes variations de niveau en fonction du temps. À plus haute fréquence, l'incertitude inhérente à la réponse du microphone en fonction de l'angle, de la courbe de réponse de la salle et de l'absorption de la salle contribue à réduire la précision. D'autre part, les bruits ou tonalités à bande étroite subissent une trop grande influence de la réponse des haut-

⁷ Schroeder, Manfred: *The Schroeder Frequency Revisited*, JASA, vol. 99, n° 5, p. 3240-3241.

parleurs et de l'acoustique de la salle. Il a donc été établi qu'une bande de deux octaves centrée sur la fréquence de 1 kHz serait le plus utile pour les canaux principaux et qu'une bande centrée sur 40 Hz serait le plus utile pour le canal d'effets basses fréquences.

Annexe A : Intensité sonore de programme

A.1 INTRODUCTION

La présente annexe expose le contexte des algorithmes de mesure de l'intensité sonore et de la crête vraie définis dans la pratique recommandée BS.1770. Le lecteur trouvera la description détaillée de ces algorithmes dans le document de référence [3].

A.2 INTENSITÉ SONORE DES PROGRAMMES

L'intensité sonore perçue d'un signal audio, ou sonie, peut être examinée de diverses façons. Les changements d'intensité sonore de courte durée sont la manifestation de la nature fréquemment changeante du contenu audio. Les changements de moment à moment au cours d'un programme, qui communiquent des changements d'atmosphère ou de contexte, en sont un exemple. Les changements d'intensité sonore de courte durée peuvent aussi se produire au passage d'un programme à un autre. Certains de ces changements peuvent être intrinsèques, étant donné que chaque programme peut avoir été créé dans un but distinct. Ils peuvent aussi provenir du fait que les programmes ont été créés suivant des méthodes de production différentes.

L'intensité sonore d'un signal audio peut aussi être examinée sur une plus grande échelle. Pour la distribution d'un programme, il est souvent nécessaire d'obtenir une valeur d'intensité sonore « composée », qui prend en compte tous les changements d'intensité sonore de courte durée intermédiaires survenant au cours de ce programme. Ce type de renseignements peut être ensuite utilisé par le radiodiffuseur pour appliquer une correction de niveau sonore globale dans un effort pour améliorer l'uniformité des niveaux d'intensité sonore dans tous les programmes et dans les passages d'un programme à un autre. Cette intensité sonore composée, dite de longue durée, peut aussi servir à comparer les divers éléments de programme, qu'ils représentent la parole, de la musique ou tout autre type de contenu audio.

En 2001, le secteur de la Radiocommunication de l'Union Internationale des Télécommunications (UIT-R) a entamé une étude visant à identifier une mesure objective de la sonie de programmes sonores types utilisés en télédiffusion. L'étude a abouti à l'adoption d'une méthode permettant de calculer l'intensité sonore de longue durée d'un signal audio. Cette mesure vise les signaux mono, stéréo et multicanaux. Des études sont toujours en cours pour bien comprendre les propriétés de l'intensité sonore de courte durée.

La première phase de l'étude de l'UIT-R a consisté à mettre au point une méthode d'essai subjective pour examiner la sonie de programmes monophoniques types. Un groupe de trois personnes formé de membres du groupe de travail 6P SRG3 de l'UIT-R a choisi 48 séquences d'essai, consistant en une vaste gamme de matériel de radiodiffusion; une séquence de référence contenant une voix féminine anglaise a été choisie pour établir une intensité sonore cible. Le niveau de lecture de la séquence de référence a été fixé à 60 dBA SPL. Outre l'utilisation d'une seule séquence de référence commune, les auditeurs de l'essai ont été encouragés à comparer toutes les séquences d'essai les unes avec les autres. Cette approche d'essai est analogue à celle

du document de référence [9], où l'auditeur est en mesure de comparer divers éléments de test entre eux durant un essai. Chacune des séquences d'essai a été répétée à deux niveaux d'amplitude, créant ainsi au total de 96 séquences audio monophoniques que les auditeurs devaient apparier. Un total de 97 auditeurs ont participé à l'essai, à cinq emplacements différents. Les résultats des essais subjectifs [10] ont constitué la base pour l'évaluation du rendement de divers algorithmes ou appareils de mesure de l'intensité sonore.

Sept entités ont présenté dix appareils/algorithmes de mesure de l'intensité sonore monophonique pour évaluation. En plus des sonomètres, deux algorithmes de mesure d'intensité sonore ont été soumis par le laboratoire d'évaluation pour servir de point de référence de rendement. Ces deux algorithmes sont de simples calculs de valeur quadratique moyenne : le premier a recours à un filtre à pondération en fréquence simple et le deuxième est sans pondération. Le filtre à pondération simple utilise une courbe de pondération B basse fréquence révisée (RLB). La comparaison des sonomètres/algorithmes de mesure de sonie soumis avec la base de données subjectives a révélé que l'algorithme à valeur quadratique moyenne pondéré en fréquence simple a présenté le meilleur rendement [11].

Pour vérifier les performances de cet algorithme, une deuxième base de données subjectives a été créée à partir d'essais subjectifs formels conduits dans l'un des emplacements d'essai. Pour cet essai, 20 auditeurs ont évalué la sonie des 96 séquences audio monophoniques suivant la même méthode subjective que la première ronde d'essais. L'analyse des résultats enregistrés dans cette deuxième base de données subjectives a confirmé les performances de l'algorithme des valeurs moyennes pondérées en fréquence.

Dans le but d'étendre l'application de l'algorithme aux signaux audio multicanaux, une troisième série d'essais a été menée, portant sur 144 séquences audio (48 séquences monophoniques, 48 stéréo et 48 multicanaux) présentées à 20 auditeurs. La même méthode subjective a été employée que dans les deux premiers essais. La séquence de référence était une voix féminine anglaise accompagnée d'une ambiance stéréo sur fond musical de faible niveau sonore. Les haut-parleurs étaient disposés suivant la configuration décrite dans la Recommandation UIT-R BS.775 [12]. Les résultats de ce troisième essai subjectif [13] ont abouti à la conception de l'algorithme de mesure de l'intensité sonore multicanal représenté à la figure A.1.

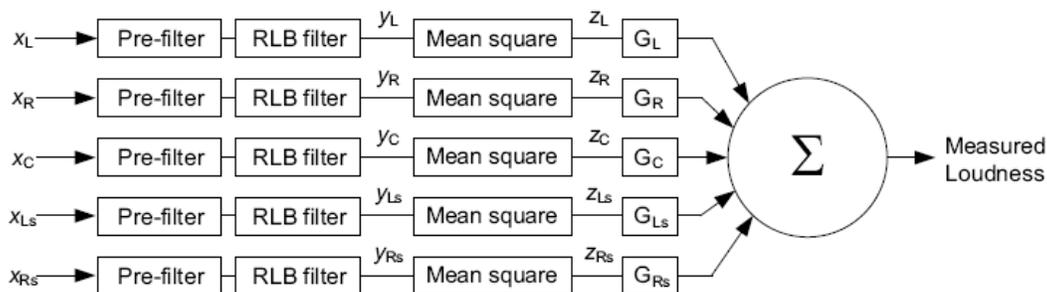


Figure A.1 Schéma fonctionnel de l'algorithme de mesure de l'intensité sonore multicanal

Pre-filter	Préfiltre
RLB	Filtre RLB
Mean square	Valeur quadratique moyenne

Measured Loudness	Intensité sonore
-------------------	------------------

L'intensité sonore de chaque canal est mesurée de manière indépendante et les valeurs sont additionnées pour donner l'intensité sonore résultante. Outre la courbe de pondération RLB, représentée à la figure A.2, un préfiltre est ajouté pour tenir compte des effets de la tête humaine. La réponse en fréquence de ce préfiltre est représentée à la figure A.3.

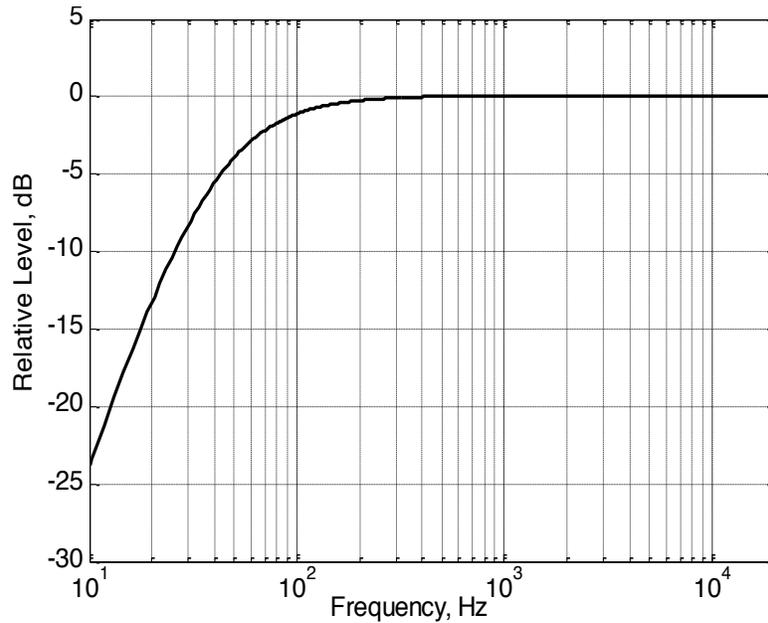


Figure A.2 Courbe de pondération RLB

Relative Level, dB	Niveau relatif (dB)
Frequency, Hz	Fréquence (Hz)

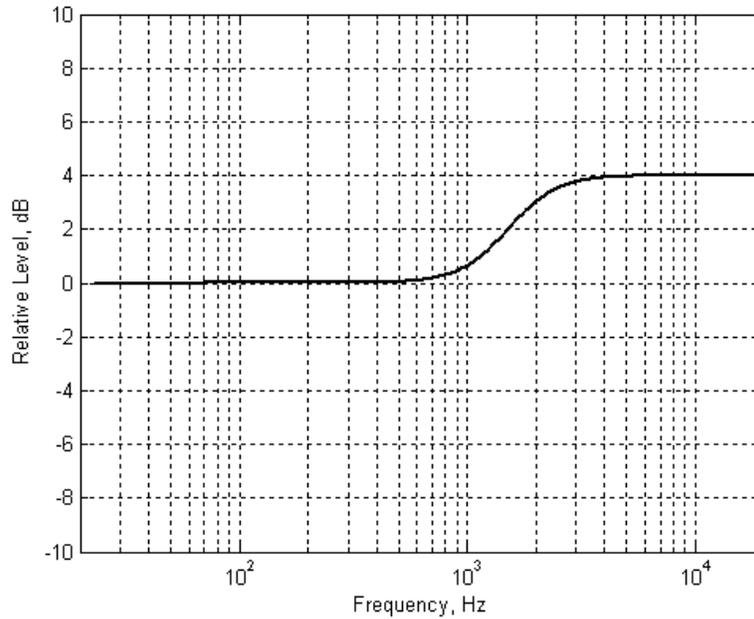


Figure A.3 Réponse du préfiltre utilisé pour tenir compte des effets acoustiques de la tête

Relative Level, dB	Niveau relatif (dB)
Frequency, Hz	Fréquence (Hz)

Les performances de l'algorithme peuvent être représentées en comparant les évaluations subjectives de l'intensité sonore aux valeurs mesurées. Cette représentation fait l'objet de la figure A.4, où l'intensité sonore mesurée de chacune des 240 séquences audio présentées au cours des trois essais subjectifs est reportée sur une courbe avec les valeurs de sonie. S'il y avait correspondance parfaite, les points se trouveraient tous sur la diagonale du diagramme. Dans ce diagramme, la corrélation entre les évaluations subjectives et les valeurs mesurées est de 0,977.

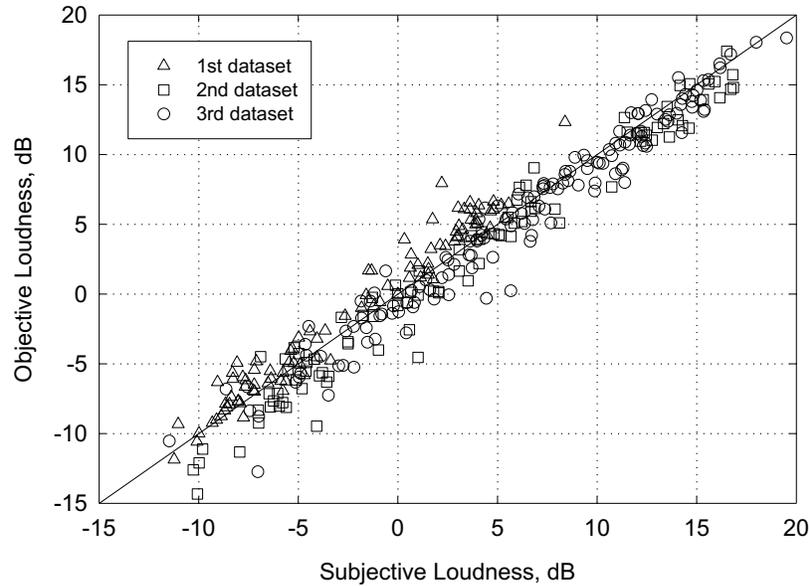


Figure A.4 Résultats regroupés pour les trois ensembles de données ($r = 0,977$).

Objective Loudness, dB	Intensité sonore objective (dB)
Subjective Loudness, dB	Intensité sonore subjective (dB)
1st dataset	Premier ensemble de données
2nd dataset	Deuxième ensemble de données
3rd dataset	Troisième ensemble de données

A.3 CRÊTE VRAIE

Les systèmes audio numériques ont considérablement simplifié la manipulation et la distribution des signaux audio. La mesure de la crête de ces signaux prend généralement la forme d'une représentation de la valeur absolue maximale d'un échantillon sur une période de mesure donnée. Cet accent mis sur les valeurs d'échantillons de crête a favorisé l'absence d'intérêt pour la forme d'onde continue sous-jacente. Cette « omission » peut conduire à des surcharges audio imprévues, à des indications de crête contradictoires et à d'autres problèmes dissimulés. Les mesures de niveau de crête vrai donnent une description plus précise du signal audio, qui peut être utile pour prévenir ces problèmes.

La figure A.5 représente la possibilité d'indications imprécises quand on utilise des crêtes-mètres d'échantillons types. Dans ce diagramme, le niveau maximum de la forme d'onde continue dépasse la valeur d'échantillon maximale. Cette sous-indication du niveau audio est faible en basses fréquences, mais peut être importante dans les fréquences plus hautes ou pour les signaux contenant des transitoires abrupts. Des incohérences dans les indications de crête d'échantillons peuvent se produire lorsque les échantillons ne tombent pas aux mêmes endroits dans le signal audio. Dans le diagramme, un léger décalage de phase dans le signal audio peut produire des indications de crête d'échantillon différentes.

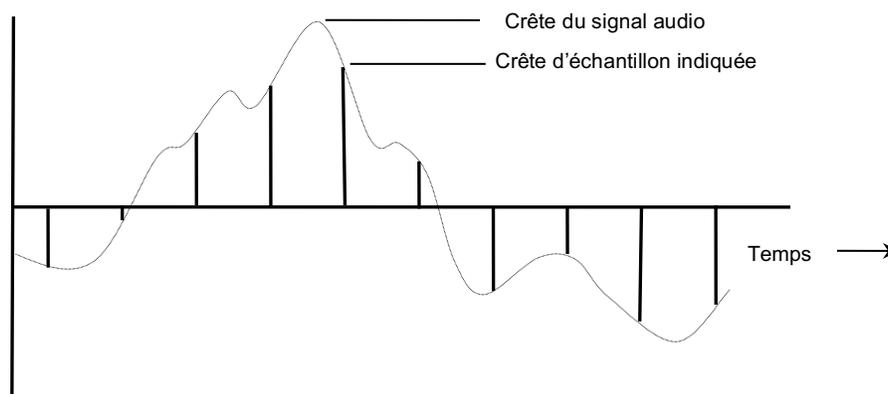


Figure A.5 Niveau de crête de signal continu et de crête d'échantillon

L'annexe 2 de la recommandation UIT-R BS.1770-1 décrit un algorithme permettant d'estimer le niveau de crête vrai dans un signal audionumérique MIC linéaire. Une description simplifiée de l'algorithme est représentée à la figure A6.

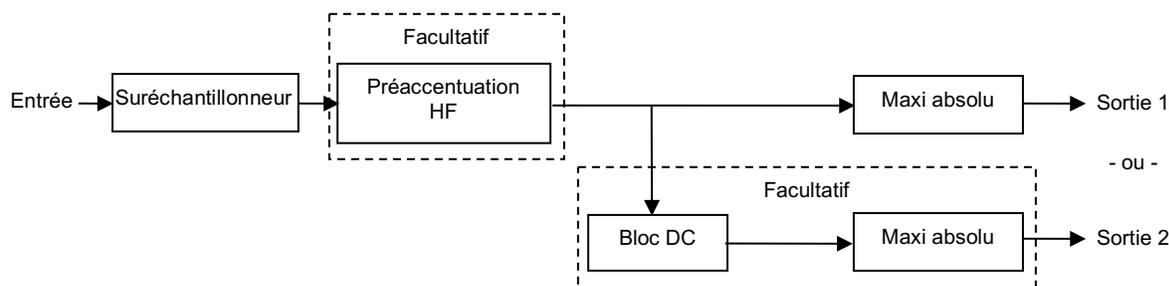


Figure A.6 Structure de base pour l'algorithme d'estimation de la crête vraie défini dans UIT-R BS.1770-1

Les signaux entrants sont suréchantillonnés à 192 kHz pour obtenir une représentation plus précise du signal audio. Cela signifie un quadruple suréchantillonnage pour les signaux échantillonnés à 48 kHz. Étant donné que la possibilité d'une sous-lecture des niveaux de crête d'échantillon est plus élevée dans les hautes fréquences, un filtre de préaccentuation facultatif peut être ajouté pour tenir compte de l'effet que produisent sur les hautes fréquences du signal audio les filtres de protection contre le repliement de spectre que l'on retrouve fréquemment dans l'équipement, dans toute la chaîne de radiodiffusion.

Les signaux audio comportant une composante DC peuvent avoir une incidence sur les niveaux de crête, les élevant ou les abaissant suivant la nature du signal. Un deuxième élément facultatif de l'algorithme de mesure de crête vraie supprime la composante DC, permettant à l'opérateur d'évaluer les niveaux tels qu'ils seraient dans un dispositif qui bloque la composante DC.

Annexe B :

Acoustique de salle et disposition des haut-parleurs

La présente PR n'a pas pour objet de discuter en détail de l'acoustique des salles et de la disposition des haut-parleurs, mais la qualité du contrôle dans les régies peut être énormément améliorée par l'observation de quelques principes fondamentaux.

Contrôler les modes propres en basse fréquence des salles en utilisant des dimensions irrégulières et des matériaux d'absorption de basse fréquence efficaces

Chaque salle a son « mode propre de résonance », déterminé par la géométrie, qui entraîne la résonance du son à certaines fréquences. Éliminer le parallélisme des murs n'élimine pas ce phénomène. Dans les petites pièces utilisées pour le contrôle, la fréquence de résonance la plus basse se situe dans la gamme audible et les crêtes de résonance sont très écartées. Si le phénomène n'est pas contrôlé, il produit des anomalies importantes dans la réponse en fréquence, qui varie de canal en canal, quelle que soit la qualité des haut-parleurs. L'égalisation seule n'est pas très utile pour résoudre ce problème, car les crêtes de résonance sont très étroites. Dans ces conditions, le mixage est difficile, chaque note d'un mélange musical pouvant ressortir et être beaucoup trop forte, alors que les notes adjacentes ne le sont pas.

Dans les salles contenant peu de gros objets, choisir les trois dimensions inégales en évitant les rapports simples entre les dimensions d'au moins ± 5 pour cent. Toujours prévoir l'absorption des basses fréquences pour amortir la réponse de la salle dans les fréquences de résonance. Les basses fréquences ne peuvent être absorbées que par des matériaux de qualité supérieure d'au moins 2 po d'épaisseur ou par un revêtement de 2 po couvrant un creux d'air plus profond. La pose de tapis sur les murs ou les planchers ne donne pratiquement rien. Les plafonds suspendus réalisés au moyen de carreau absorbant haute performance à revêtement intérieur de matériau absorbant de 4 po d'épaisseur peuvent absorber efficacement les basses fréquences pour amortir le mode propre de résonance de la salle dans un plan. L'absorption de coins profonds, soit à l'angle mur-mur ou mur-plafond, peut aussi être efficace.

Distribuer les matériaux d'absorption acoustique et diffusifs de qualité supérieure de manière aléatoire et éviter les surfaces parallèles dures

Dans une salle où la réverbération est excessive, il est difficile d'entendre le détail d'un mixage. De plus, l'effet d'oscillations parasites, où une série de fréquences moyennes à hautes résonnent, se produit si deux surfaces dures sont parallèles, particulièrement si les autres surfaces de la salle sont absorbantes.

Pour réduire les deux effets, répartir aléatoirement sur les murs et les plafonds des matériaux d'absorption de qualité supérieure de 2 po d'épaisseur, de manière à éliminer les surfaces parallèles où deux zones opposées sont réfléchissantes. Réduire au minimum l'utilisation de tapis muraux et d'autres matériaux absorbants minces, qui ne sont utiles qu'à très haute fréquence. S'il est impossible de rendre les surfaces absorbantes, dans le cas des fenêtres, par exemple, l'inclinaison de la vitre intérieure peut éliminer le parallélisme des surfaces. Les éléments

diffusifs sont aussi utiles pour produire un espace ayant une réverbération raisonnable, mais contrôlée.

Placer les haut-parleurs et l'absorption de manière à empêcher les réflexions discrètes

Une surface dure plane située derrière la position de mixage crée un effet de « filtre-peigne » provenant des haut-parleurs de contrôle principaux. La nature du filtre-peigne change en fréquence avec le déplacement du technicien de mixage vers l'avant ou l'arrière de la salle. Un effet analogue, à fréquences plus hautes, se produit lorsque le son direct du haut-parleur est réfléchi sur une grande face de pupitre. L'effet sur le technicien de mixage peut dissimuler les problèmes réels de filtre-peigne dans le mixage.

Les réflexions discrètes sur le mur arrière devraient être éliminées par la pose d'une combinaison de matériau d'absorption et de diffusion sur cette surface. Songer à placer les haut-parleurs de manière à empêcher les réflexions contre la face du pupitre d'atteindre les oreilles du technicien de mixage.

Réduire la résonance de la gamme médiane/basse par le positionnement et l'égalisation

Un haut-parleur à réponse parfaitement uniforme soumis à une mesure anéchoïque peut avoir un son très différent dans la gamme basse-médiane dans des installations, suivant sa position dans la salle, en raison du renforcement et de la suppression entraînés par les réflexions provenant de surfaces réfléchissantes adjacentes à proximité du haut-parleur.

On peut atténuer cet effet en évitant de placer les haut-parleurs à égale distance de deux murs, en éloignant les haut-parleurs des surfaces réfléchissantes ou en posant un matériau absorbant les fréquences moyennes sur les murs adjacents. Étant donné que les effets sont plutôt larges, l'égalisation peut être très utile pour uniformiser la réponse.

Choisir des haut-parleurs ayant une réponse uniforme et une large dispersion acoustique

La mesure ci-dessus ne peut corriger la situation lorsque les haut-parleurs ont une réponse en fréquence irrégulière ou si la réponse varie énormément suivant l'angle d'écoute. On ne peut compter sur de tels haut-parleurs pour représenter la balance des fréquences du mixage final.

Choisir des haut-parleurs qui ont un son uniforme quand on se déplace de la position centrale d'écoute dans l'axe du haut-parleur jusqu'au bord de la zone d'écoute. Rechercher la preuve que la réponse en fréquence des haut-parleurs est uniforme sur toute la gamme audio. Songer à utiliser un caisson d'extrêmes graves pour étendre la réponse en basse fréquence.

Pour obtenir de plus amples renseignements, consulter les documents de référence *Audio Monitoring in Contemporary Post-Production Environments* [16] et *Surround Sound: Up and Running* » [17].

Annexe C : Correction acoustique

La correction acoustique des salles est nécessaire dans les environnements de contrôle pour s'attaquer à de nombreux problèmes survenant en raison de l'interaction du son produit par les haut-parleurs et de la salle. Les effets négatifs de l'acoustique d'une petite salle produisent plusieurs phénomènes secondaires audibles, dont la distorsion d'imagerie due aux réflexions non voulues et aux anomalies de réponse en fréquence, qui sont particulièrement graves dans les basses fréquences.

Traditionnellement, l'égalisation des salles était exécutée au moyen de systèmes analyseurs/égaliseurs ayant recours à la mesure d'un bruit rose au moyen d'un microphone et d'égaliseurs paramétriques ou graphiques. Ces systèmes présentent de sérieuses limites, pour les raisons suivantes :

- 1) Une seule mesure dans une salle ne peut fournir suffisamment de renseignements sur les performances à basse fréquence dans la zone d'écoute.
- 2) Le bruit rose ne peut servir à mesurer que la réponse en amplitude et ne fournit aucune information sur la réponse dans le temps.
- 3) Les égaliseurs paramétriques ou graphiques n'ont pas assez de résolution, même avec 30 bandes de fréquences.
- 4) Les filtres IIR utilisés dans ces égaliseurs peuvent souffrir d'anomalies de phases, particulièrement lorsqu'ils deviennent plus étroits.

Pour être efficace, la correction acoustique des salles doit :

- Saisir des données du domaine temporel pour pouvoir bien prendre en compte les effets de la réflexion.
- Saisir des données du domaine fréquentiel avec une résolution suffisamment élevée dans les basses fréquences pour s'attaquer aux problèmes types des petits locaux.
- Combiner plusieurs mesures prises dans la zone d'écoute pour tenir compte des variations de basse fréquence entraînées par les ondes stationnaires.
- Réduire la réverbération dans les basses fréquences (*modal ringing*).

Les exigences relatives à la correction acoustique efficace des salles peuvent être remplies par l'utilisation de filtres FIR dans la solution d'égalisation. Toutefois, les approches FIR classiques standard ne suffisent pas. Des méthodes comprenant la pondération fréquentielle sont typiquement employées dans la conception des filtres FIR pour distribuer la puissance des filtres de façon non linéaire avec les fréquences, permettant d'obtenir davantage de puissance de correction aux basses fréquences. De plus, la moyenne spatiale de plusieurs mesures ne suffit pas non plus. La distribution des problèmes acoustiques dans l'espace n'est pas uniforme et certains lieux présentent des problèmes plus graves que d'autres. Il faut donc aussi employer une méthode de pondération spatiale non-linéaire quand on combine les mesures.

Annexe D :

Aide-mémoire de configuration de contrôle de référence pour la télévision

Pour les espaces de postproduction :

1. Copier la première tonalité, onde sinusoïdale 1 kHz à -20 dB FS, dans un poste de travail audio numérique, puis la copier dans chacun des canaux à employer, p. ex. 5.1.
2. Lire le fichier en réglant les commandes de niveau de sortie du poste de travail au gain unitaire, l'entrée pupitre au gain unitaire et la commande de niveau général du pupitre sur le gain unitaire. Pour le moment, on peut laisser la commande de niveau de contrôle sur un réglage bas.
3. Régler le niveau de chaque canal, tour à tour, sur le niveau d'essai d'onde sinusoïdale pour l'appareil de mesure utilisé, p. ex. sur -20 dB FS pour les appareils numériques ou sur 0 VU pour les vumètres.
4. Copier la deuxième tonalité d'essai, le bruit rose à bande limitée à -20 dB FS, dans chaque canal du poste de travail audio numérique.
5. Lire un canal à la fois.
6. Placer un sonomètre au point correspondant au centre de la tête de l'opérateur en position normale, l'orientant vers le canal de haut-parleur utilisé. Garder le corps perpendiculaire au sonomètre, d'un côté de l'appareil, car le niveau peut être modifié par une forte réflexion contre le corps.
7. Régler la commande de niveau de contrôle général sur une position normalisée, reproductible, qui sera utilisée au mixage.
8. Régler les commandes de niveau de canal individuel, comme les commandes de gain d'un amplificateur de puissance ou les commandes de gain des haut-parleurs amplifiés, sur le niveau de pression acoustique (SPL) mesuré à pondération C et mesure lente indiqué dans le tableau D.1.

Tableau D.1 Niveau de pression acoustique de référence

Catégories	Volume de la salle (pi ³)	SPL (dB relatifs à 20 µN/m ²)
I, II (Salles de mixage)	> 20,000	85*
	10,000 < 19,999	82
	5,000 < 9,999	80
	1,500 < 4,999	78
	< 1,499	76
III (Salles de montage parfois utilisées pour le mixage)	Est fonction de l'utilisation de la pièce. Pour le montage, doit être commandé par le monteur pour utilisation avec le matériel dont il dispose. Pour le mixage final de programme, suivre les recommandations visant les catégories I et II ci-dessus.	
IV (Cabines, camions)	< 1,500	76
V (Casques d'écoute)	Utiliser un coupleur 2 cm ³ et régler le niveau du signal 400 Hz à 78 dB.	
* Conformément à SMPTE RP 200 [6]		

Première tonalité d'essai, onde sinusoïdale :

http://www.atsc.org/refs/a85/440Hz_left_Ch-20dB.wav

Deuxième tonalité d'essai, bruit rose à bande limitée :

http://www.atsc.org/refs/a85/MidRngPinkNoise_-20dB.wav

Aide-mémoire très bref :

S'il est impossible d'utiliser le bruit rose à bande limitée pour étalonner les niveaux de contrôle, on peut employer l'échantillon de parole mentionné dans la section 10.4 pour régler le niveau de contrôle de façon appropriée. Veiller à ce que le trajet du signal de l'appareil de lecture émettant l'échantillon de parole au système de contrôle soit réglé sur le gain unitaire. Monter le fichier source mono dans le canal simple pour la lecture mono, dans les canaux gauche et droit à -3 dB chacun pour la lecture stéréo et dans le canal central des systèmes à 5.1 canaux et autres systèmes multicanaux. Faire jouer l'échantillon de parole et régler le gain de contrôle général de manière à obtenir le meilleur confort d'écoute de la voix. Étant donné que l'intensité sonore de l'échantillon de parole est de -24 LKFS, les programmes mixés de manière à ce que l'élément d'ancrage du programme tel que le dialogue corresponde à ce niveau, auront approximativement la même sonie.

http://www.atsc.org/refs/a85/Speech_sample.wav

Annexe E : Plage d'intensité sonore

La zone de confort est la plage d'intensité sonore dans laquelle un auditeur accepte les changements de sonie à l'intérieur d'éléments de contenu et entre des éléments. Une expérience subjective a été conduite pour établir cette plage ainsi que les autres points de « tolérance à la sonie ». Les résultats de l'expérience sont indiqués dans la figure E1.

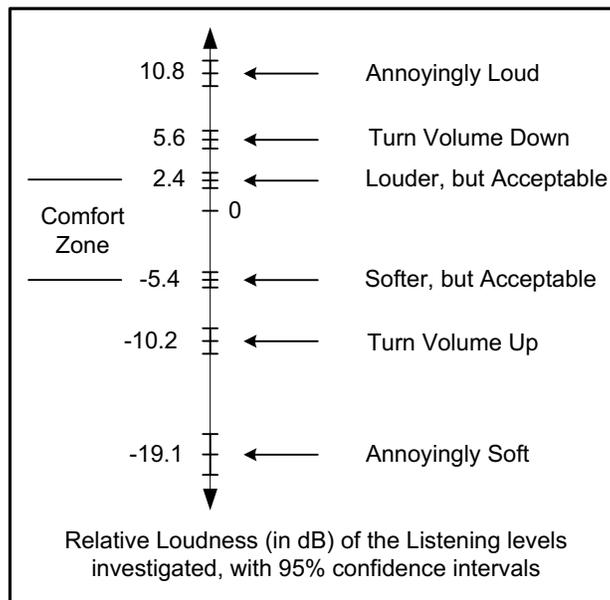


Figure E.1 Seuils d'intensité sonore critiques

Annoyingly Loud	Trop fort, contrariant
Turn Volume Down	Baisser le volume
Louder, but Acceptable	Fort, mais acceptable
Softer, but Acceptable	Faible, mais acceptable
Turn Volume Up	Monter le volume
Annoyingly Soft	Trop faible, contrariant
Comfort Zone	Zone de confort
Relative Loudness (in dB) of the Listening levels investigated, with 95% confidence intervals	Intensité sonore relative (dB) des niveaux d'écoute étudiés, avec intervalles de confiance de 95 %

L'expérience reproduisait les passages entre les contenus de longue durée et de courte durée d'une même chaîne et d'une chaîne à une autre. Les auditeurs étaient placés dans une situation d'écoute ou de visionnement type et ils étaient invités à passer d'un échantillon de contenu à un autre, de longue et de courte durée. On leur avait présenté cinq paires d'éléments monophoniques

composées d'un élément de référence et d'un élément d'essai reproduits par un seul haut-parleur situé devant eux.

Les auditeurs devaient régler le niveau de lecture général jusqu'à ce que l'élément de référence soit reproduit au niveau sonore qu'ils considèrent comme un « volume confortable », le point zéro sur le schéma. L'expérimentateur leur aura demandé ensuite de régler la commande de volume d'essai (le terme « volume » est plus familier qu'intensité sonore pour la plupart des auditeurs) sur l'un des six niveaux représentés sur le schéma. Les auditeurs pouvaient écouter les éléments d'essai et de référence à tour de rôle à volonté. Une fois qu'ils avaient décidé du niveau où se trouvait la différence d'intensité sonore demandée, le décalage était enregistré. Les questions étaient posées de façon aléatoire et l'ordre de présentation des paires d'éléments de référence et d'essai était aussi aléatoire d'un auditeur à l'autre. Les éléments de référence et d'essai provenaient d'une autre expérience qui fournissait une sonothèque de contenus de sonie constante.

Étant donné que les éléments de référence et de test étaient réputés d'égale intensité sonore, le décalage de gain que les auditeurs ont appliqué à l'élément de test en réponse aux questions de l'expérimentateur constitue une mesure directe de la zone de confort des auditeurs et des autres seuils d'intensité sonore critiques étudiés.

Il est intéressant de noter qu'une augmentation de gain de deux ou trois dB dans le niveau sonore suffit à faire passer la sonie d'un programme type de la zone de confort de l'auditeur vers le point où il aimerait baisser le volume. Il y a plus d'écart du côté bas du point de « volume confortable » (représenté par « 0 » ici).

Le niveau de bruit ambiant dans la salle d'écoute où ont eu lieu les essais était assez faible; analogue à celui du salon d'une maison située en campagne, par une soirée tranquille. Étant donné que l'on peut raisonnablement croire que le point « trop faible, contrariant » se situe quelque part au-dessus du niveau de bruit ambiant de l'environnement d'écoute, le chiffre de -19,1 dB est probablement étroitement fonction du niveau du bruit ambiant. Les autres points sont suffisamment écartés du niveau ambiant pour que leur répartition relative ne soit pas touchée.

Annexe F –

Détails sur le contrôle de gamme dynamique AC-3

F.1 APERÇU GÉNÉRAL DU DRC

Le système AC-3 N'A PAS pour but de reproduire le traitement audio utilisé dans le système de télévision analogique courant. Il vise plutôt 1) à fournir une intensité sonore de dialogue uniforme entre les programmes et 2) à permettre aux auditeurs de réduire les excursions de la dynamique autour de ce niveau sonore commun s'ils le souhaitent.

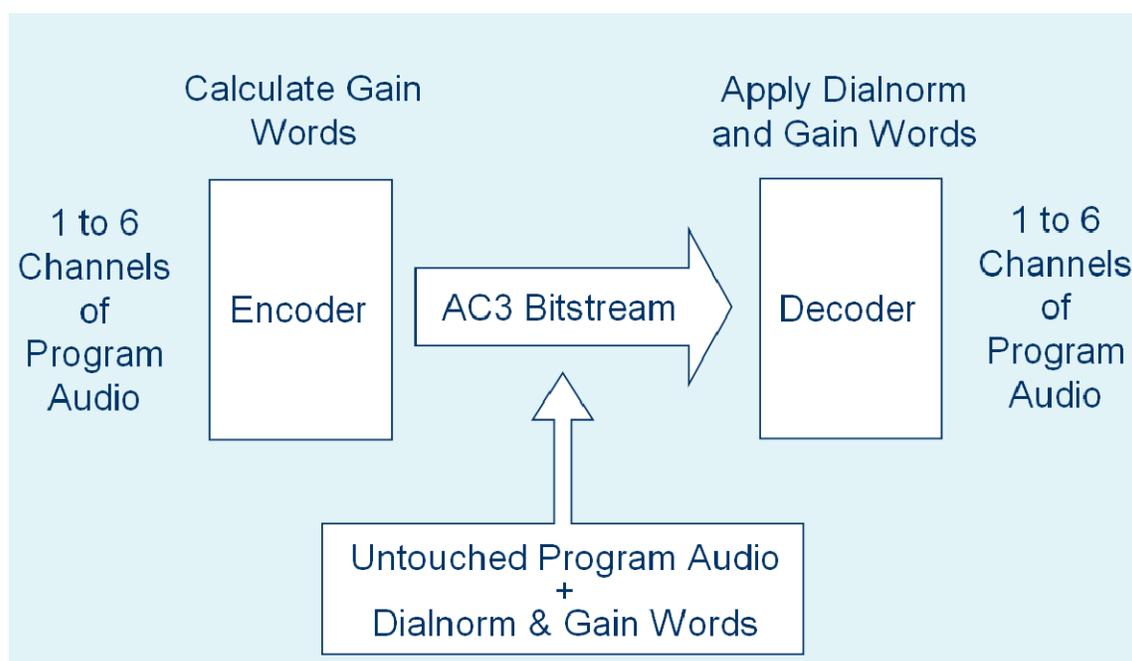


Figure F.1 Principe de base du contrôle de gamme dynamique (DRC) AC-3

Calculate Gain Words	Calcul des mots de gain
Apply Dialnorm and Gain Words	Application de Dialnorm et des mots de gain
1 to 6 Channels of Program Audio	1 à 6 canaux d'audio de programme
Encoder	Encodeur
AC3 Bitstream	Flux binaire AC3
Decoder	Décodeur
1 to 6 Channels of Program Audio	1 à 6 canaux d'audio de programme
Untouched Program Audio + Dialnorm & Gain Words	Audio de programme original + Dialnorm et mots de gain

Le principe de base de la partie contrôle de gamme dynamique du système AC3 est d'exécuter tous les calculs nécessaires à la réduction de la dynamique du programme pour qu'il

corresponde au « profil » de dynamique sélectionné dans l'encodeur AC-3, puis d'envoyer la version originale (gamme dynamique étendue) du programme (contenu) au décodeur AC-3 avec les mots de gains DRC décrivant les modifications de gain nécessaires pour réduire la dynamique du programme pour qu'elle corresponde au profil sélectionné (voir figure F.1). L'auditeur a ensuite la possibilité d'appliquer les mots de gain, ou non, suivant ses exigences d'écoute. Un auditeur peut choisir de ne pas appliquer les mots de gain, auquel cas il entend la dynamique originale du programme, et un autre auditeur peut choisir d'écouter le programme dans une gamme dynamique réduite. Dans les deux cas, l'intensité sonore du dialogue, ou « point d'ancrage », de tous les programmes est normalisée, que les données DRC soient appliquées ou non.

La sélection du « profil » de compression revient au producteur du programme, car il lui permet de faire un choix artistique en fonction duquel la dynamique du contenu est réduite. Le système AC-3 peut fournir des profils distincts pour deux modes (mode ligne et mode RF), comme il est expliqué plus loin.

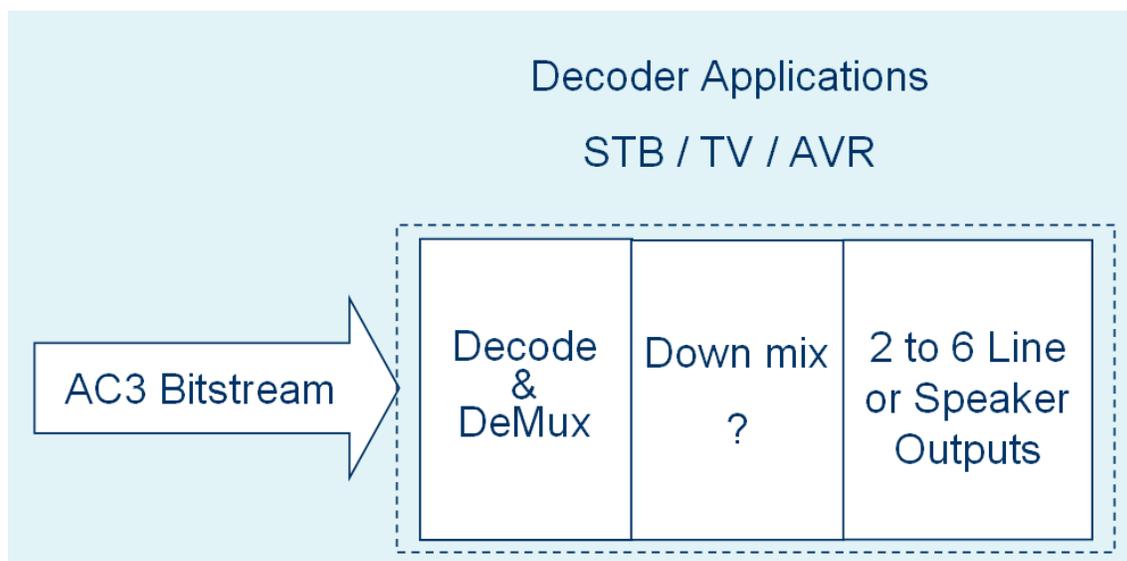


Figure F.2 Exigences relatives au décodeur AC-3 (mode ligne)

Decoder Applications	Applications de décodeur
STB / TV / AVR	Boîte numérique/TV/RAV
AC3 Bitstream	Flux binaire AC3
Decode & DeMux	Décodage et démultiplexage
Down mix ?	Mixage réducteur
2 to 6 Line or Speaker Outputs	2 à 6 lignes ou sorties haut-parleurs

Quand des décodeurs sont utilisés, deux « modes » principaux sont nécessaires. Le premier, représenté à la figure F2, constitue une source de signaux de niveau ligne servant à fournir les signaux audio destinés aux cinémas-maison et autres applications haute fidélité. Dans ce cas, seule une légère compression est nécessaire, voire aucune.

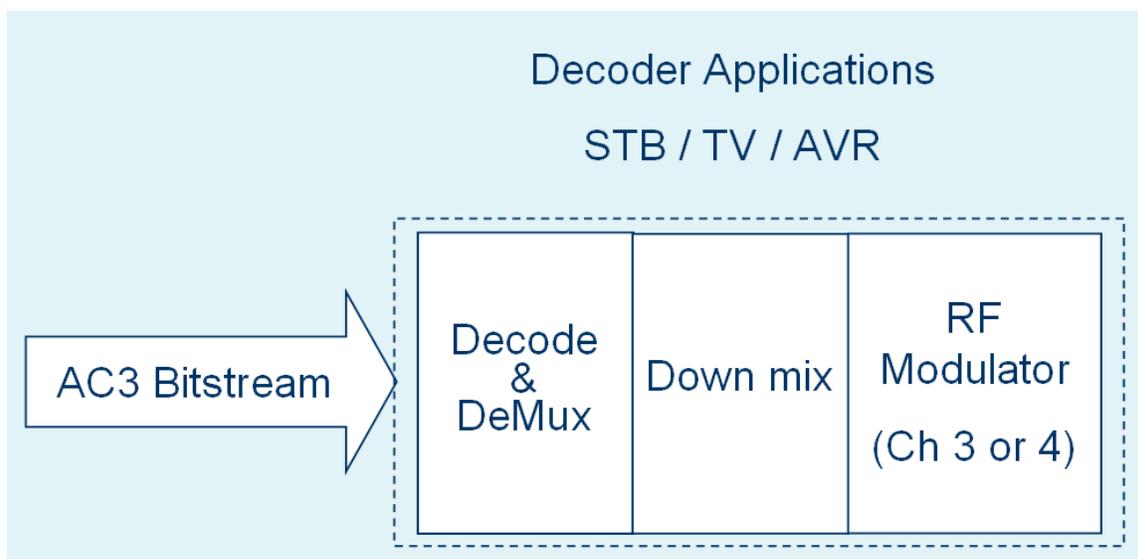


Figure F.3 Exigences relatives au décodeur AC-3 (mode RF).

Decoder Applications	Applications de décodeur
STB / TV / AVR	Boîte numérique/TV/RAV
AC3 Bitstream	Flux binaire AC3
Decode & DeMux	Décodage et démultiplexage
Down mix	Mixage réducteur
RF Modulator (Ch 3 or 4)	Modulateur RF (chaîne 3 ou 4)

Le deuxième mode (figure F.3) est utilisé pour alimenter un modulateur RF. Dans ce cas, le programme doit subir une réduction par mixage et l'intensité sonore doit être accentuée pour assurer une correspondance raisonnable avec les autres signaux analogiques hertziens utilisés par le téléviseur. Le système doit aussi réduire davantage la gamme dynamique que dans le cas précédent, en raison du niveau de signal programme amplifié, pour la faire correspondre à la gamme dynamique réduite des émissions de télévision existantes typiques, et possiblement en raison des conditions d'écoute.

F.2 PLACEMENT DES MÉTADONNÉES DANS LES FLUX BINAIRES DRC AC-3

Le système AC-3 fournit des instructions de gain pour les deux types de réduction de dynamique ainsi qu'une fonction de limitation au cas où l'augmentation de niveau inhérente au mixage réducteur serait suffisante pour produire l'écrtage.

Les mots de gain du mode dit de ligne représenté à la figure F.2 sont appelés mots « dynrng ». Ceux qui sont destinés au mode dit RF représenté à la figure F.3 sont appelés mots « compr ».

Le flux binaire AC-3 transporte les deux types de mots de gain, dynrng et compr (comme l'indique la figure F.4) ainsi que les mots de gain de limitation, au besoin. Les décodeurs AC-3 peuvent appliquer le contrôle de dynamique RF ou de ligne, suivant la préférence de l'auditeur (bien que certains fabricants d'équipement grand public limitent les options dans certains appareils).

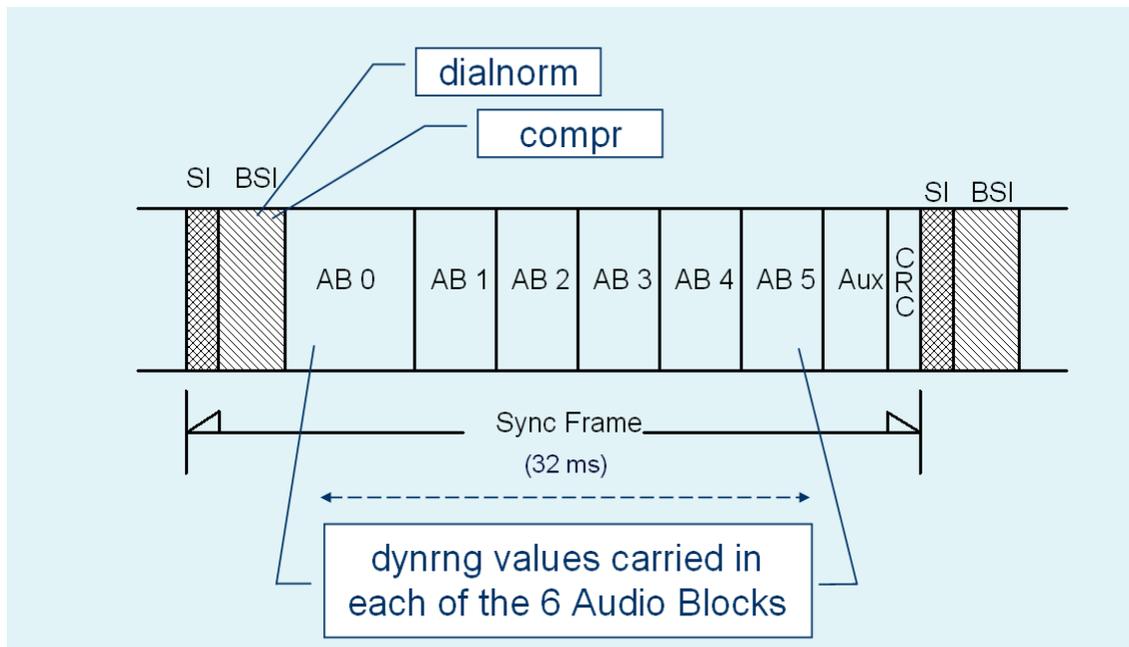


Figure F.4 Placement des métadonnées de DRC AC-3 dans le flux binaire

SI	SI
BSI	BSI
dialnorm	dialnorm
compr	compr
Aux	Aux.
CRC	CRC
Sync Frame (32 ms)	Synchro trame (32 ms)
dynrng values carried in each of the 6 Audio Blocks	Valeurs dynrng transportées dans chacun des 6 blocs audio (AB)

Tous les mots de contrôle de gamme dynamique, y compris les instructions de limite de gain, sont générés dans l’encodeur AC-3, envoyés au décodeur AC-3 et appliqués dans le décodeur AC-3. Les données transmises comprennent les données dialnorm, qui sont utilisées pour normaliser l’intensité sonore du dialogue ou « point d’ancrage » de chaque programme dans son ensemble.

Les mots de gain dialnorm et compr sont envoyés au décodeur à intervalles de 32 ms, tandis que les mots de gain dynrng sont transportés à six fois ce rythme, c.-à-d. environ toutes les 5,3 ms.

F.3 CALCULS DES MOTS DE GAIN AC-3

Comme le représente la figure F.5, la première tâche consiste à calculer l’intensité sonore et les niveaux de crête normalisés du contenu du programme (d’autres détails sont indiqués à la figure F.6). Le paramètre acomod indique à l’encodeur le nombre de canaux que contient le programme.

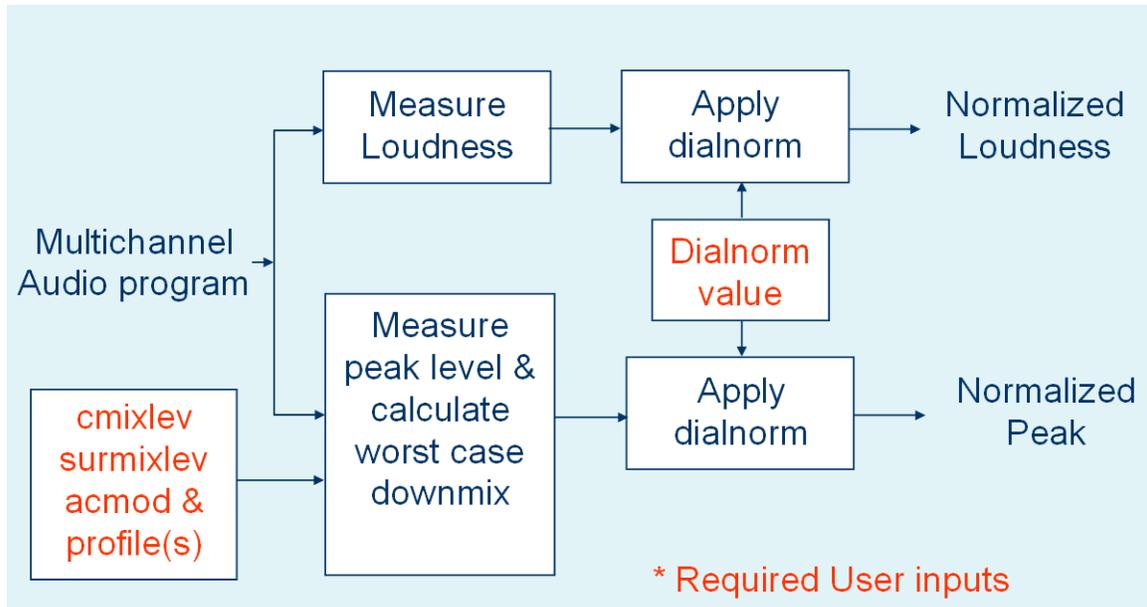


Figure F.5 Calculs de mots de gain AC-3 par l'encodeur (partie 1).

Measure Loudness	Mesure de l'intensité sonore
Apply dialnorm	Application de dialnorm
Normalized Loudness	Intensité sonore normalisée
Multichannel Audio program	Programme audio multicanal
Dialnorm value	Valeur dialnorm
cmixlev surmixlev acmod & profile(s)	cmixlev, surmixlev, acmod et profil(s)
Measure peak level & calculate worst case downmix	Mesure du niveau de crête et calcul du pire cas de mixage réducteur
Normalized Peak	Crête normalisée
* Required User inputs	*Intervention de l'utilisateur nécessaire

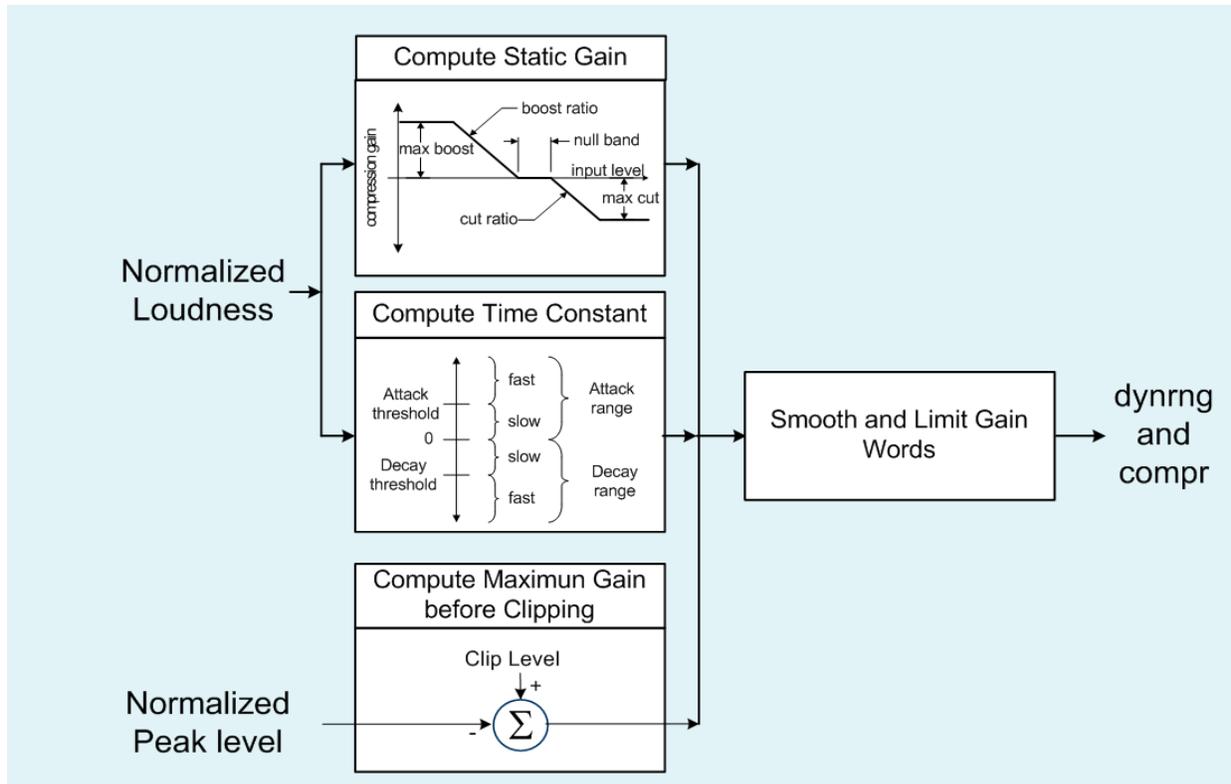


Figure F.6 Calculs de mots de gain AC-3 par l'encodeur (partie 2)

Compute Static Gain	Calcul du gain statique
compression gain	gain de compression
max boost	amplification maxi
boost ratio	taux d'amplification
null band	bande sans correction
input level	niveau d'entrée
max cut	atténuation maxi
cut ratio	taux d'atténuation
Normalized Loudness	Intensité sonore normalisée
Compute Time Constant	Constante de temps de calcul
Attack threshold	Seuil d'attaque
Decay threshold	Seuil de chute
fast	rapide
slow	lent
Attack range	Gamme d'attaque
Decay range	Gamme de chute
Smooth and Limit Gain Words	Adoucir et limiter les mots de gain
dynrng and compr	dynrng et compr
Compute Maximum Gain before Clipping	Calcul du gain maximum avant écrêtage
Clip Level	Niveau de saturation
Normalized Peak level	Niveau de crête normalisé

Pour les mots de gain *dynrng*, l'intensité sonore est calculée pour chaque sous-ensemble de 256 échantillons (les blocs audio représentent 512 échantillons – 5,3 ms – d'audio, mais en raison de la structure de blocs de « chevauchement-addition », l'intensité sonore de chaque bloc est représentée par la moyenne de deux mesures de 256 échantillons). Les valeurs d'intensité sonore sont donc produites à intervalles de ~5,3 ms.

L'intensité sonore de chaque canal est calculée individuellement, puis les résultats sont additionnés pour obtenir l'intensité sonore globale du programme. La courbe de pondération de l'intensité sonore est représentée à la figure F.8.

Le calcul du niveau de crête doit tenir compte des coefficients de réduction par mixage du canal du centre et des canaux d'ambiance (*cmixlev* et *surmixlev*) sélectionnés (idéalement) durant le processus de production. La sortie du calcul est le niveau de crête maximum du pire cas de réduction par mixage.

Les valeurs d'intensité sonore et de crête sont normalisées au moyen de la valeur *dialnorm* de programme avant d'être passées aux étages de calcul suivants.

Les données d'intensité sonore et de niveau de crête utilisées pour les calculs de mots de gain « *compr* » représentent ces valeurs sur toute l'étendue d'une synchro trame AC-3 (32 ms).

Le décodeur audio comprend une fonction de chevauchement-addition, ce qui permet d'appliquer les mots de gain *compr* en douceur même lorsqu'ils sont synchronisés abruptement.

La partie DRC du système AC-3 doit traiter des contenus de programme de tous genres, c'est pourquoi elle comprend une méthode d'ajustement des constantes de temps d'attaque et de chute déterminées automatiquement suivant le contenu du programme.

Les données audio après normalisation de l'intensité sonore sont mises en correspondance dans le profil DRC sélectionné pour générer une valeur de gain statique. Elles sont aussi utilisées pour sélectionner l'une des quatre constantes de temps, selon que l'intensité sonore d'entrée est plus élevée (attaque) ou plus basse (chute) que l'intensité sonore ajustée, et de combien.

Si l'intensité sonore d'entrée est plus élevée que l'intensité sonore ajustée et que la différence est plus grande que le seuil d'attaque, alors la constante de temps choisie est l'attaque rapide (*fast attack*). Il s'agit généralement d'une constante de temps très rapide, conçu pour assurer la convergence rapide du compresseur pour les événements très forts.

Si l'intensité sonore d'entrée est plus élevée que l'intensité sonore ajustée, mais d'une valeur ne dépassant pas le seuil d'attaque, alors la constante de temps choisie est l'attaque lente (*slow attack*). Il s'agit généralement d'une constante de temps modérément rapide, conçue pour fournir une convergence assez rapide du contrôleur de dynamique pour les événements modérément forts.

Si l'intensité sonore d'entrée est plus basse que l'intensité sonore ajustée, mais d'une valeur ne dépassant pas le seuil de chute, alors la constante de temps choisie est la chute lente (*slow decay*). Il s'agit généralement d'une constante de temps très lente, conçue pour fournir un relâchement lent et donc inaudible du compresseur durant les passages calmes.

Si l'intensité sonore d'entrée est plus basse que l'intensité sonore ajustée et que la différence est plus grande que le seuil de chute, alors la constante de temps choisie est la chute rapide (*fast decay*). Il s'agit généralement d'une constante de temps modérée, conçue pour fournir un relâchement rapide à la fin d'événements très forts.

L'autre principale partie du processus de génération des données DRC est la fonction de limitation de crête. Le niveau de saturation est connu (0 dB FS), ainsi la différence entre ce niveau et le niveau de crête normalisé du programme est le gain maximum admissible avant

l'écrtage. Si les mots de gain statique appellent davantage de gain, le bloc d'adoucissement et de limite restreint les mots de gain statique au gain maximum admissible. La constante de temps qui vient d'être calculée fait en sorte que les changements de gain ne seront pas suffisamment abrupts pour produire des phénomènes parasites désagréables et qu'ils sont appropriés pour le contenu du programme.

Il est à noter que deux processus s'exécutent en parallèle : l'un au débit des blocs audio (~5,3 ms) et générant les mots de gain *dynrng* et l'autre au débit de la synchro trame (32 ms), générant les mots de gain *compr*.

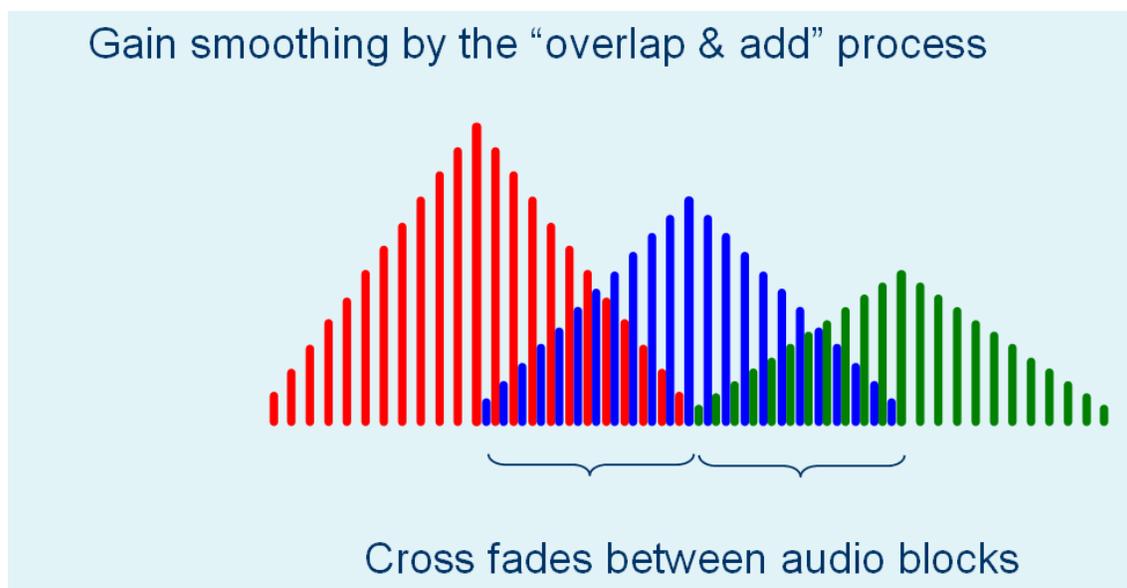


Figure F.7 Adoucissement du gain

Gain smoothing by the "overlap & add" process	Adoucissement du gain par processus de chevauchement-addition
Cross fades between audio blocks	Enchaînés entre les blocs audio

Il s'agit d'une très simple représentation de la façon dont les transitions de « chevauchement-addition » entre les blocs audio se produisent et de l'effet qu'elles ont sur les pas de gain (quantifié en pas de 0,2 dB pour *dynrng* et en pas de 0,4 dB pour *compr*) appliqués par le système de contrôle de gamme dynamique.

La forme triangulaire présentée à la figure F.7 est une approximation (pour faciliter le dessin) de la fonction de fenêtrage réelle appliquée aux données.

Les mots *dynrng* s'appliquent à un groupe entier de 512 échantillons, mais chacun de ses groupes chevauche de moitié le groupe adjacent, de sorte que l'effet produit lorsque les valeurs des échantillons sont additionnées est un enchaîné entre les groupes, qui adoucit les pas de gain qui les séparent.

Les mots *compr* s'appliquent à une trame de synchronisation entière, mais les trames de synchronisation sont aussi soumises au processus de chevauchement-addition, ce qui a pour effet d'adoucir aussi les pas de gain *compr*.

F.4 PONDÉRATION D'INTENSITÉ SONORE ET COURBES D'ENTRÉES-SORTIE DE DRC

La courbe de pondération DRC accentue davantage les basses fréquences que la courbe de pondération B, qui est normalement utilisée pour pondérer les sons de « niveau modéré » (c.-à-d. près des niveaux d'écoute de la télévision). (Voir figure F.8.)

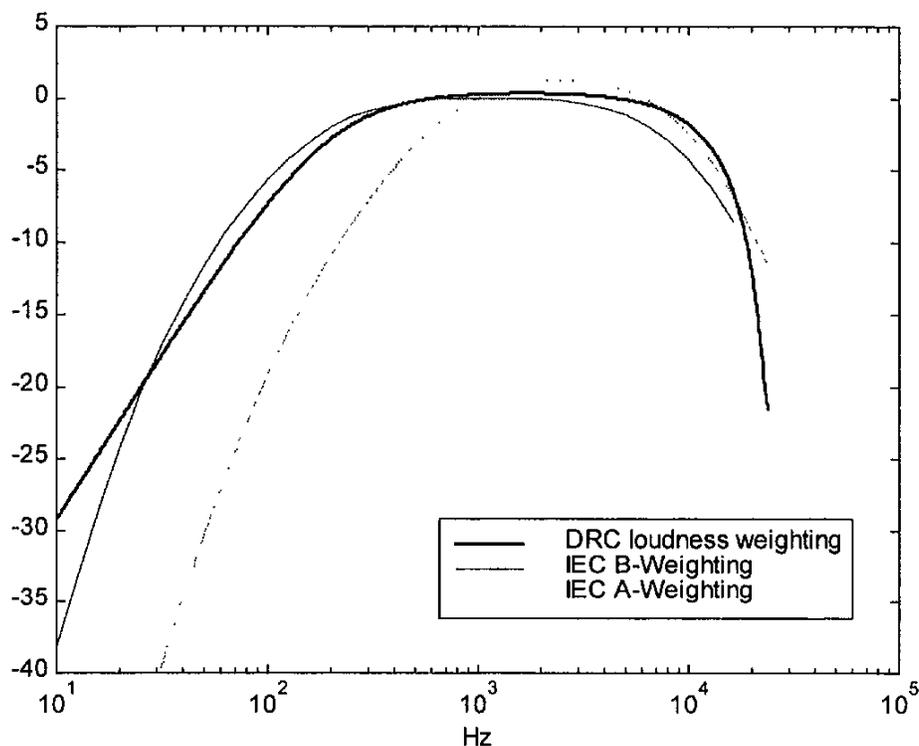


Figure F.8 Courbe de pondération d'intensité sonore DRC

DRC loudness weighting	Pondération d'intensité sonore DRC
IEC B-Weighting	Pondération B de l'IEC
IEC A-Weighting	Pondération A de l'IEC

L'intensité sonore totale d'une source multicanal est calculée comme la somme des niveaux de puissance des valeurs quadratiques moyennes pondérées de chaque canal. La pondération et l'addition des valeurs d'intensité sonore sont des calculs antérieurs aux mesures d'intensité sonore de l'UIT, qui peuvent être considérées comme un perfectionnement de cette méthode.

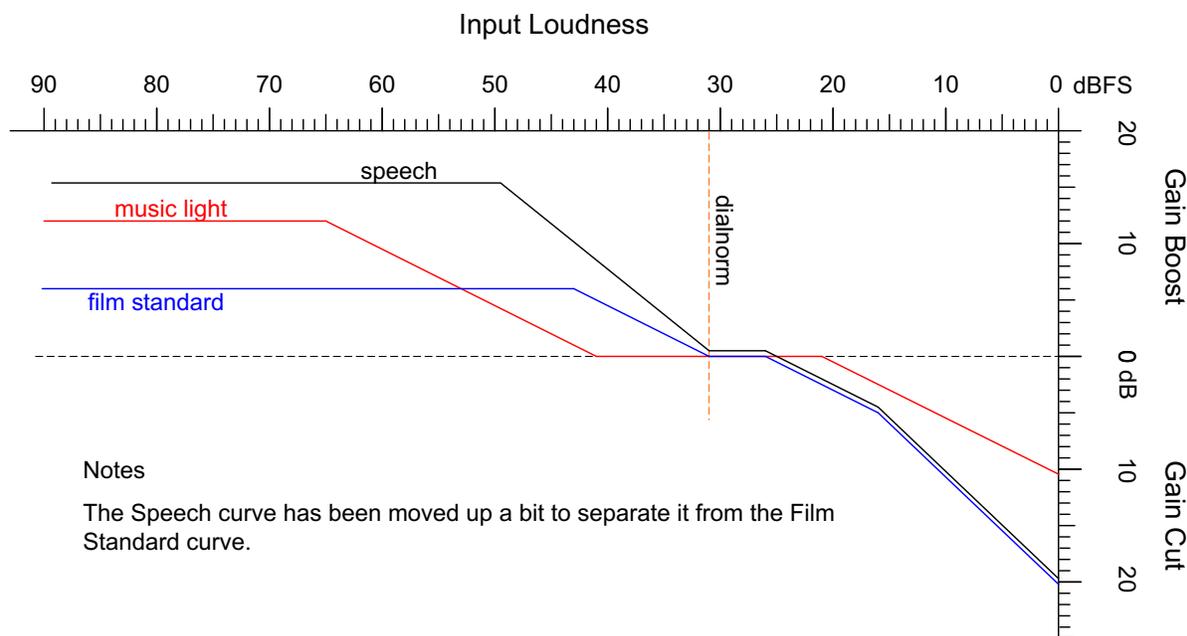


Figure F.9 Caractéristiques d'entrée-sortie des profils DRC

Input Loudness	Intensité sonore d'entrée
dBFS	dB FS
Gain Boost	Amplification de gain
Gain Cut	Atténuation de gain
Notes The Speech curve has been moved up a bit to separate it from the Film Standard curve.	Notes La courbe du profil de parole (Speech) a été déplacée légèrement vers le haut pour la distinguer de la courbe du profil Film standard
speech	parole
music light	musique légère
film standard	film standard
dialnorm	dialnorm
0 dB	0 dB

Les zones linéaires (sans correction) des profils Film standard et Speech ne sont pas symétriques autour de la valeur dialnorm, parce que dans les contenus de programmes types et avec le degré de réduction de gamme dynamique voulu, l'attaque de zone d'amplification relativement rapide conjuguée avec un temps de chute plus lent a tendance à laisser le programme dans la zone linéaire asymétrique (voir figure F.9).

F.5 PARAMÈTRES D'ENCODEUR POUR LE RÉGLAGE DES MÉTADONNÉES DRC

Voici quelques-unes des options de menu que l'on retrouve au panneau avant de l'encodeur AC-3 Dolby 569. Il existe un logiciel de commande à distance compatible PC pour le Dolby 569, qui présente tous les paramètres sur un même écran. L'appareil de création et de visualisation de métadonnées Dolby 570 a des menus de panneau avant analogues ainsi qu'un logiciel de commande à distance (beaucoup plus facile à utiliser).

Parmi les paramètres de réglage signalons:

Menu principal *Setup* -> Audio Service -> Channel Mode -> Choisir une des options *1/0*, *2/0*, *3/2*, etc.

Menu principal *Setup* -> Audio Service -> Dialog Level -> Régler sur *-1 dB to -31 dB*.

Menu principal *Setup* -> BSI Parameters -> Center Downmix level -> Choisir une des options *-3.0 dB*, *-4.5 dB*, *-6.0 dB*.

Menu principal *Setup* -> BSI Parameters -> Surround Downmix level -> Choisir une des options *-3.0 dB*, *-4.5 dB*, *-6.0 dB*.

Menu principal *Setup* -> Dynamic Range -> Line Mode -> Choisir une des options : *Speech*, *Music Light*, *Music Standard*, *Film Light*, *Film Standard*, *Disabled*.

Menu principal *Setup* -> Dynamic Range -> RF Mode -> Choisir une des options : *Speech*, *Music Light*, *Music Standard*, *Film Light*, *Film Standard*, *Disabled*.

F.6 COMMANDES EXÉCUTABLES AU DÉCODEUR DRC

Il est à noter que :

- Il y a peu de terminologie commune entre les appareils grand public pour les modes DRC de ligne, RF ou hors fonction.
- Dans le décodeur type, le mode ligne est appelé « *Normal* », le mode RF, « *Late Night* » (fin de soirée) et l'absence de compression de dynamique, « *Theatre* » (cinéma).
- En règle générale, le consommateur peut choisir un de ces trois modes.
- Dans certains décodeurs (haut de gamme), l'auditeur peut régler le pourcentage de compression appliquée dans les modes ligne et RF.

Annexe G : Paramètres de données AC-3

Les valeurs par défaut définies dans certains encodeurs AC-3 peuvent n'avoir aucun rapport avec le contenu particulier de l'opérateur; on ne devrait donc pas s'y fier. Le fabricant de l'encodeur AC-3 devrait être consulté pour obtenir des conseils.

Les principaux paramètres de données AC-3 sont présentés dans les tableaux G.1 à G.4.

Tableau G.1 Métadonnées de contrôle essentielles

Fonction	Variable dans le flux	Description
<i>Dialogue Level</i>	dialnorm	(Niveau de dialogue) Commande la normalisation du dialogue.
<i>Channel Mode</i>	acmod	(Mode de canal) Indique le nombre et le type de canaux actifs.

Tableau G.2 Métadonnées de contrôle – encodeur

Fonction	Description
<i>Bitrate</i>	(Débit binaire) Débit binaire codé.
<i>RF Overmodulation Protection</i>	(Protection contre la surmodulation RF) Si elle est activée, l'encodeur AC-3 inclut la préaccentuation dans ses calculs de compression en mode RF.
<i>DC Filter</i>	(Filtre DC) Applique un filtre passe-haut 3 Hz de blocage DC avant l'encodage AC-3.
<i>Lowpass Filter</i>	(Filtre passe-bas) Applique un filtre passe-bas aux canaux d'entrée principaux avant l'encodage AC-3.
<i>LFE Lowpass Filter</i>	(Filtre passe-bas LFE) Applique un filtre passe-bas 120 Hz au canal LFE avant l'encodage AC-3.
<i>Surround 3 dB Attenuation</i>	(Atténuation ambiance 3 dB) Atténue les canaux ambiance de 3 dB avant l'encodage.
<i>Surround Phase Shift</i>	(Déphasage ambiance) Avant l'encodage, crée les canaux d'ambiance déphasés nécessaires pour créer une sortie Lt/Rt dans un décodeur capable de décoder en L, C, R, S au moyen de Dolby Pro Logic .

Tableau G.3 Métadonnées de contrôle de base – décodeur

Fonction	Description
<i>LFE Channel</i>	(Canal LFE) Indique si le canal LFE est présent. Disponible seulement dans les modes de canaux 3/2, 3/1, 3/0, 2/2 et 2/1.
<i>DRC Line Mode Profile</i>	(Profil DRC de mode ligne) Indique la configuration de compression de DRC prédéfinie pour le décodage du mode ligne.
<i>DRC RF Mode Profile</i>	(Profil DRC de mode RF) Indique la configuration de compression de DRC prédéfinie pour le décodage du mode RF.
<i>Dolby Surround Mode</i>	(Mode Dolby Surround) Indique si le flux binaire codé à deux canaux contient un programme Dolby Surround (Lt/Rt) et doit être décodé en Dolby Pro Logic.
<i>Dolby Surround EX Mode*</i>	(Mode Dolby Surround EX) Indique si l'audio est encodé en tant que flux Dolby Surround EX et nécessite le décodage Surround EX.

<i>Preferred Stereo Downmix*</i>	(Réduction stéréo préférée) Indique la réduction par mixage préférée : Lt/Rt (codé Pro Logic) ou Lo/Ro (stéréo seulement).
<i>Center Downmix Level</i>	(Niveau de réduction centre) Indique le niveau de réduction visant le canal C lorsque l'utilisateur final n'a pas de haut-parleur central.
<i>Lt/Rt C Downmix Level*</i>	(Niveau de réduction C Lt/Rt) Lorsque la réduction stéréo est Lt/Rt, indique le niveau de réduction visant le canal C lorsque l'utilisateur final n'a pas de haut-parleur central.
<i>Lo/Ro C Downmix Level*</i>	(Niveau de réduction C Lo/Ro) Lorsque la réduction stéréo est Lo/Ro, indique le niveau de réduction visant le canal C lorsque l'utilisateur final n'a pas de haut-parleur central.
<i>Surround Downmix Level</i>	(Niveau de réduction ambiance) Indique le niveau de réduction visant les canaux d'ambiance lorsque l'utilisateur final n'a pas de haut-parleurs d'ambiance. Il est à noter que le réglage 0 (-999 dB) élimine les canaux d'ambiance.
<i>Lt/Rt S Downmix Level*</i>	(Niveau de réduction ambiance Lt/Rt) Lorsque la réduction stéréo est Lt/Rt, indique le niveau de réduction visant les canaux d'ambiance lorsque l'utilisateur final n'a pas de haut-parleurs d'ambiance.
<i>Lo/Ro S Downmix Level*</i>	(Niveau de réduction ambiance Lo/Ro) Lorsque la réduction stéréo est Lo/Ro, indique le niveau de réduction visant les canaux d'ambiance lorsque l'utilisateur final n'a pas de haut-parleurs d'ambiance.
*Paramètres de flux binaire étendu. Non pris en charge par tous les décodeurs.	

Tableau G.4 Métadonnées informatives

Fonction	Description
<i>Audio Production Info</i>	(Info de production audio) Indique si les valeurs des paramètres Mixing Level (niveau de mixage) et Room Type (type de salle) sont transportés dans le flux binaire.
<i>Bitstream Mode</i>	(Mode de flux) Décrit le service audio transporté dans le flux binaire.
<i>Copyright</i>	(Droit d'auteur) Indique si le flux binaire codé est protégé par droit d'auteur.
<i>Mixing Level</i>	(Niveau de mixage) Indique le niveau de pression acoustique absolu du canal de dialogue principal durant la session de mixage final.
<i>Room Type</i>	(Type de salle) Indique la taille et l'étalonnage de la salle de mixage utilisée pour le mixage final.
<i>Original Bitstream</i>	(Flux original) Indique si le flux binaire AC-3 codé est la version maîtresse.
<i>A/D Converter Type*</i>	(Type de convertisseur A-N) Indique le réglage du type de convertisseur analogique-numérique.

Annexe H :

Aide-mémoire à l'intention des techniciens de station et de MVPD

Gestion de l'intensité sonore

L'aide-mémoire qui suit doit être proposé dans un petit document distinct.

H.1 INTRODUCTION

Note : Le présent aide-mémoire est fondé sur la Pratique recommandée par l'ATSC A/85 : *Techniques d'établissement et de maintien de l'intensité sonore pour la télévision numérique*, (la PR) et le texte complet est reproduit dans l'annexe H de la PR.

Cet aide-mémoire résume les recommandations exposées dans la PR et fournit des conseils aux radiodiffuseurs et autres distributeurs de programmes télévisuels pour le contrôle et le maintien d'une intensité sonore uniforme dans leurs stations et chaînes de télévision.

H.2 PORTÉE

Cet aide-mémoire n'est pas destiné à remplacer la PR complète. Sa portée est limitée à servir de guide du type « comment faire pour... » à l'intention des opérateurs de station de télévision et des MVPD. Le lecteur est encouragé à étudier la PR en entier pour obtenir des renseignements plus détaillés et le contexte du présent guide. En cas de contradiction entre l'aide-mémoire et la PR, cette dernière a préséance sur le présent aide-mémoire.

Cet aide-mémoire est fondé sur l'utilisation d'un système de métadonnées statiques⁸ (voir la section **Error! Reference source not found.** de la PR).

H.3 DÉFINITIONS

BS.1770 – Officiellement UIT-R BS.1770 [3]. Cette recommandation définit un algorithme qui permet d'obtenir une valeur numérique indiquant la sonie du contenu mesuré. Les sonomètres et les outils de mesure dans lesquels est mis en œuvre l'algorithme BS.1770 indiquent l'intensité sonore en « LKFS ».

Contenu (*content*) – Matériel ou substance du produit distribué par un diffuseur.

dialnorm – Paramètre de métadonnées AC-3, de valeur numérique égale à la valeur absolue du Niveau de dialogue, transporté dans le flux binaire AC-3. C'est un code de 5 bits non signé, qui indique l'écart négatif du Niveau de dialogue moyen par rapport à 0 LKFS. Les valeurs valides vont de 1 à 31. La valeur 0 est réservée. Les valeurs 1 à 31 sont interprétées -1 à -31. Le décodeur applique une valeur de réduction de gain égale à la différence entre -31 et la valeur dialnorm.

⁸ Les métadonnées dynamiques constituent une solution de rechange au système de métadonnées statiques (fixes). Consulter la section 7.5 pour obtenir plus de détails sur l'utilisation des métadonnées dynamiques.

Élément d'ancrage (*Anchor Element*) – Point de référence de la sonie (intensité sonore perçue) ou élément autour duquel d'autres éléments sont équilibrés pour produire le mélange final de contenu ou qui importe le plus au téléspectateur raisonnable quand il règle la commande de volume.

LKFS – Intensité sonore (Loudness) avec pondération K, par rapport au niveau sonore maximum (*Full Scale*), mesurée au moyen d'un appareil qui met en œuvre l'algorithme spécifié par l'UIT-R BS.1770 [3]. Une unité LKFS est équivalente à un décibel.

Niveau de dialogue (*Dialog Level*) – L'intensité sonore de l'Élément d'ancrage, exprimée en LKFS⁹.

Intensité sonore cible (*Target Loudness*) – Valeur spécifiée pour l'Élément d'ancrage (c.-à-d. Niveau de dialogue), établie pour faciliter l'échange de contenu entre un fournisseur et un opérateur.

H.4 GESTION DE L'INTENSITÉ SONORE

Point à retenir : Le but est de présenter au téléspectateur un contenu audio d'intensité sonore uniforme dans les publicités, les émissions et les changements de chaîne.

H.5 EXIGENCE DE LA FCC

Le document ATSC A/53, partie 5:2010 [1] rend obligatoire le transport du paramètre dialnorm et de valeurs dialnorm réglées de façon appropriée.

Point à retenir : Régler le paramètre dialnorm de l'encodeur AC-3 de la station pour qu'il corresponde à l'intensité sonore du niveau de dialogue moyen du contenu.

H.6 MESURE DU CONTENU DISTRIBUÉ

Voir la section 5 de la PR.

H.6.1 Contenu longue durée

Une partie représentative du contenu dominée par le dialogue type (c.-à-d. non par des cris ou des chuchotements) devrait être isolée et mesurée. En l'absence de dialogue, l'intensité sonore de l'élément du contenu sur lequel se fonderait un téléspectateur raisonnable pour régler le volume devrait être mesurée. Si aucune des deux techniques n'est possible ou pratique, l'intensité sonore de tout le contenu devrait être mesurée. Si le contenu contient des périodes importantes de silence, consulter la section 5.2 de la PR.

Point à retenir : Mesurer l'intensité sonore du contenu longue durée dans les passages où il y a des dialogues. Il s'agit de la valeur du Niveau de dialogue du contenu. Le

⁹ Le terme « Niveau de dialogue » est fondé sur l'utilisation répandue du dialogue comme ancrage pour le mixage du contenu; historiquement, on a considéré le dialogue comme l'Élément d'ancrage de la plupart des programmes.

Niveau de dialogue (en unités LKFS) devrait correspondre à la valeur dialnorm de l'encodeur AC-3.

H.6.2 Contenu courte durée

Voir la section 5.2.4 de la PR.

Point à retenir : Mesurer l'intensité sonore de tous les canaux audio¹⁰ et de tous les éléments intégrés de la piste sonore, sur toute la durée du contenu de courte durée. La valeur de la mesure de l'intensité sonore (en unités LKFS) devrait correspondre à la valeur dialnorm de l'encodeur AC-3.

H.6.3 Bulletins de nouvelles ou autre programmation en direct

Le principe de mesurer l'intensité sonore du dialogue du contenu s'applique aux productions en direct réalisées en temps réel au fur et à mesure que la production progresse.

L'objet des mesures d'intensité sonore exécutées durant un événement en direct est de guider le mixeur pour qu'il puisse produire le contenu à une intensité sonore correspondant au réglage dialnorm de l'encodeur AC-3 de la station.

Un sonomètre BS.1770 peut être utile dans les mixages effectués dans un milieu bruyant ou lorsqu'il est impossible de maintenir un niveau de contrôle uniforme. Voir la section 5.2.2 de la PR.

Point à retenir : Utiliser un sonomètre BS.1770 pour faciliter l'alignement de l'intensité sonore du contenu en temps réel sur le réglage dialnorm de l'encodeur AC-3.

H.6.4 Contenu enregistré dans des fichiers

Le stockage dans des fichiers facilite l'automatisation des mesures d'intensité sonore et l'ajustement de l'intensité sonore du contenu ou de la valeur dialnorm pouvant avoir été attribuée au contenu. Voir la section 5.2.5 de la PR.

Point à retenir : Veiller à ce que le contenu enregistré dans un fichier corresponde au réglage dialnorm de la station, défini dans l'encodeur AC-3.

H.7 INTENSITÉ SONORE CIBLE POUR FACILITER L'ÉCHANGE DE PROGRAMMES

Voir la section 6 de la PR.

L'Intensité sonore cible est une valeur spécifiée pour le Niveau de dialogue et établie pour faciliter l'échange de contenu entre un fournisseur et un opérateur.

Pour la distribution ou l'échange de contenu sans métadonnées (et en l'absence d'accord préalable entre les parties en matière d'intensité sonore), l'ATSC définit une valeur d'Intensité sonore cible de -24 LKFS, qui sert à établir un niveau de fonctionnement commun à employer

¹⁰ Le LFE n'est pas compris.

avec cette valeur fixe de dialnorm. De légers écarts de mesure par rapport à cette valeur, de l'ordre de ± 2 dB, sont prévus en raison de l'incertitude des mesures, et ils sont acceptables. Le niveau d'intensité sonore cible des contenus ne devrait pas être fixé au-dessus ou au-dessous de cette plage.

Point à retenir : Pour le contenu sans métadonnées, utiliser la valeur d'Intensité sonore cible de -24 LKFS.

H.8 MÉTHODES DE CONTRÔLE EFFICACE DE L'INTENSITÉ SONORE ENTRE PROGRAMME ET CONTENU INTERSTITIEL

Une grande variation d'intensité sonore durant les transitions peut être gérée avec efficacité en adoptant les pratiques ci-dessous.

Pour les opérateurs qui utilisent un système dialnorm statique, voir la section **Error! Reference source not found.** de la PR :

- a) Veiller à ce que tout le contenu réponde à l'Intensité sonore cible et que la valeur dialnorm corresponde à ce niveau.
- b) Employer un dispositif de commande de gain de fichiers pour faire correspondre à la valeur cible le Niveau de dialogue des contenus non conformes.
- c) Employer un appareil de traitement de l'intensité sonore en temps réel pour faire correspondre à la valeur cible le Niveau de dialogue des contenus non conformes.

Point à retenir : Veiller à ce que tous les contenus audio de programmes et de messages publicitaires correspondent à la valeur dialnorm de l'encodeur AC-3. Utiliser un sonomètre BS.1770 pour vérifier le Niveau de dialogue du contenu audio.

H.9 RÉGLAGE dialnorm PAR UNE STATION AFFILIÉE

Voir la section 7.3.5 de la PR.

Un opérateur (station, station affiliée, MVPD, etc.) recevant un contenu livré avec une intensité sonore fixe et qui ne fera l'objet d'aucun traitement ou ajustement de gain après le récepteur, devrait régler la valeur de dialnorm dans son encodeur AC-3 pour qu'elle corresponde au Niveau de dialogue spécifié par le créateur dans le réseau. Si un gain ou une perte fixe est appliqué dans la chaîne de signaux, la valeur dialnorm de l'encodeur AC-3 devrait être décalée en conséquence par rapport au Niveau de dialogue indiqué par le créateur.

Si un traitement d'intensité sonore est appliqué au contenu audio du créateur, la valeur de l'Intensité sonore cible du processeur devrait correspondre à la valeur dialnorm de l'encodeur AC-3. Pour un complément d'information sur le traitement audio, consulter la section 9.3.

Point à retenir : Régler la valeur dialnorm de l'encodeur AC-3 sur le Niveau de dialogue indiqué par le créateur média (avec ajustement).

H.10 INSERTION DE CONTENU PAR LA STATION DE TÉLÉVISION OU LE MVPD

Dans le cas d'insertion de messages publicitaires ou de segments locaux par la station de télévision ou le MVPD, l'opérateur devrait veiller à ce que le Niveau de dialogue de l'insertion locale coïncide avec le réglage dialnorm du flux audio inséré.

Point à retenir : Veiller à ce que le Niveau de dialogue du contenu inséré corresponde au réglage dialnorm du flux audio inséré.

Si le produit du créateur destiné au réseau est décodé en bande de base, il faut mesurer l'intensité sonore du flux audio décodé et régler la valeur dialnorm du ré-encodeur AC-3 pour qu'elle corresponde à l'intensité sonore mesurée, pour l'étape d'encodage suivante. Dans ce cas, soit que l'opérateur modifie l'intensité sonore du contenu du créateur pour qu'elle corresponde à la valeur cible du système de l'opérateur, soit qu'il utilise la valeur d'intensité sonore du contenu du créateur (mesurée) pour régler la valeur dialnorm de l'étape suivante d'encodage AC-3. À cette étape de recodage, l'opérateur doit aussi veiller à ce que les autres paramètres de métadonnées soient réglés de façon appropriée.

Point à retenir : Si le produit du créateur destiné au réseau est décodé en bande de base, veiller à ce que le Niveau de dialogue mesuré du contenu corresponde au réglage dialnorm de l'étape suivante d'encodage AC-3.

H.11 CONTRÔLE DE GAMME DYNAMIQUE (DRC) AC-3

Le système AC-3 comprend des profils DRC pour le « mode ligne » et pour le « mode RF ». Bien que la sélection de ces paramètres puisse être utile à l'opérateur et au téléspectateur pour limiter la plage d'intensité sonore globale, on ne devrait pas compter sur le DRC pour corriger les variations d'intensité sonore entre les programmes, entre les programmes et les messages publicitaires ou entre les stations de télévision ou les chaînes câblées ainsi que dans les changements de chaînes. Voir la section 8.3 et l'annexe F de la PR.

Point à retenir : On ne devrait pas compter sur le contrôle de gamme dynamique AC-3 pour atténuer les variations d'intensité sonores se produisant entre les programmes et les messages publicitaires et au passage d'une station à une autre.

Annexe I :

Aide-mémoire à l'intention des mixeurs et des monteurs son qui créent du contenu (Messages publicitaires, matériel de promotion et programmation avec gestion de l'intensité sonore des signaux audio)

L'aide-mémoire qui suit doit être proposé dans un petit document distinct.

I.1 INTRODUCTION

Note : Le présent aide-mémoire est fondé sur la Pratique recommandée par l'ATSC A/85: *Techniques d'établissement et de maintien de l'intensité sonore pour la télévision numérique*, (la PR) et le texte complet est reproduit dans l'annexe I de la PR.

Cet aide-mémoire résume les recommandations exposées dans la PR et fournit des conseils aux mixeurs et aux monteurs son qui créent des contenus audio pour la télévision numérique. Il vise un objectif de gestion de l'intensité sonore des signaux audio cohérente avec le but artistique. Le terme « contenu » englobe messages publicitaires, matériels de promotion et programmation. Le terme « contenu interstitiel » s'applique aux messages publicitaires et au matériel de promotion.

I.2 PORTÉE

Cet aide-mémoire n'est pas destiné à remplacer la PR complète. Sa portée est limitée à servir de guide du type « comment faire pour... » à l'intention des mixeurs et monteurs son participant à la création de contenu audio pour la télévision. Le lecteur est encouragé à lire la PR pour découvrir les détails et le contexte de ce guide. En cas de contradiction entre l'aide-mémoire et la PR, cette dernière a préséance sur le présent aide-mémoire¹¹.

I.3 DÉFINITIONS

BS.1770 – Officiellement UIT-R BS.1770 [3]. Cette recommandation définit un algorithme qui permet d'obtenir une valeur numérique indiquant la sonie du contenu mesuré. Les sonomètres et les outils de mesure dans lesquels est mis en œuvre l'algorithme BS.1770 indiquent l'intensité sonore en « LKFS ».

dB TP – Décibel du niveau de crête vrai par rapport au signal numérique maximum (suivant UIT-R BS.1770, annexe 2 [3])

dialnorm – Paramètre de métadonnées AC-3, de valeur numérique égale à la valeur absolue du Niveau de dialogue, transporté dans le flux binaire AC-3. C'est un code de 5 bits non signé,

¹¹ Cet aide-mémoire est fondé sur l'utilisation d'un système de métadonnées statiques (voir la section 0 de la PR). Les métadonnées dynamiques constituent une solution de rechange au système de métadonnées statiques (fixes). Consulter la section 7.5 pour obtenir plus de détails sur l'utilisation des métadonnées dynamiques.

qui indique l'écart négatif du Niveau de dialogue moyen par rapport à 0 LKFS. Les valeurs valides vont de 1 à 31. La valeur 0 est réservée. Les valeurs 1 à 31 sont interprétées -1 à -31. Le décodeur applique une valeur de réduction de gain égale à la différence entre -31 et la valeur dialnorm.

LKFS – Intensité sonore (Loudness) avec pondération K, par rapport au niveau sonore maximum (*Full Scale*), mesurée au moyen d'un appareil qui met en œuvre l'algorithme spécifié par l'UIT-R BS.1770 [3]. Une unité LKFS est équivalente à un décibel.

Niveau de dialogue (*Dialog Level*) – L'intensité sonore de l'Élément d'ancrage, exprimée en LKFS¹².

Intensité sonore cible (*Target Loudness*) – Valeur spécifiée pour l'Élément d'ancrage (c.-à-d. Niveau de dialogue), établie pour faciliter l'échange de contenu entre un fournisseur et un opérateur.

I.4 ENVIRONNEMENT DE CONTRÔLE

Disposer d'un environnement de contrôle convenable est crucial pour l'obtention de résultats de mixage satisfaisants. Pour maîtriser ce sujet, veuillez lire les renseignements de base dans la section 10.1 ainsi que la section 10.2.

Il est essentiel d'utiliser le niveau de contrôle approprié, qui varie suivant le volume de la pièce. Pour un complément d'information, consulter la section 10.4 de la PR et le tableau 10.2, ci-dessous :

Tableau 10.2 Niveau de pression acoustique de référence

Catégories	Volume du local en pieds cubes	Niveau de pression acoustique en dB relatifs à 20 µN/m ²
I, II	> 20 000	85*
	10 000 < 19 999	82
	5 000 < 9 999	80
	1 500 < 4 999	78
	< 1 499	76
III	Est fonction de l'utilisation du local. Pour le montage, peut être ajusté par le monteur selon l'utilisation du matériel dont il dispose. Pour le mixage final de programme, suivre les recommandations visant les catégories I et II ci-dessus.	
IV	< 1 500	76
V		Utiliser un coupleur 2 cm ³ et régler le niveau du signal 440 Hz à 74 dB.
* Conformément à SMPTE RP 200 [6]		

Consulter à l'annexe D la synthèse de la configuration et de l'étalonnage de la salle de mixage.

Point à retenir : Le but est de configurer l'environnement d'écoute une fois de manière appropriée, puis de veiller à toujours écouter au niveau défini, pour la création de contenu. Ceci est vrai même quand on utilise un casque d'écoute pour le contrôle.

¹² Le terme « Niveau de dialogue » est fondé sur l'utilisation répandue du dialogue comme ancrage pour le mixage du contenu; historiquement, on a considéré le dialogue comme l'Élément d'ancrage de la plupart des programmes.

I.5 CONTRÔLE DE NIVEAU UIT-R BS.1770

La Recommandation UIT-R BS.1770 fournit une nouvelle technique de mesure pour le contrôle des niveaux sonores. L'utilisation d'outils de mesure prenant en charge les méthodes de mesure BS.1770 par tous les intervenants de la production audio aidera l'industrie à gérer l'intensité sonore des signaux audio des contenus tout en respectant le but artistique. La PR contient une discussion détaillée sur la recommandation BS.1770 et sur la façon dont elle fonctionne (voir l'annexe A). Une explication plus brève de la meilleure façon d'employer la technique est fournie dans la section 5.2.

Point à retenir : Utiliser des outils de mesure conformes à la recommandation BS.1770!

I.6 INTENSITÉ SONORE DE CONTENU DURANT LE MIXAGE

Point à retenir : Le niveau de contrôle étant dûment réglé, toujours mixer en se fiant à son oreille. Utiliser un outil de contrôle d'intensité sonore BS.1770 pour confirmer ce que l'on entend.

I.7 INTENSITÉ SONORE CIBLE POUR LE CONTENU SANS MÉTADONNÉES

Pour la distribution ou l'échange de contenu sans métadonnées¹³, la valeur de l'Intensité sonore cible devrait être -24 LKFS. De légers écarts de mesure par rapport à cette valeur, de l'ordre de ± 2 dB, sont prévus en raison de l'incertitude des mesures, et ils sont acceptables. Le niveau d'intensité sonore cible des contenus ne devrait pas être fixé au-dessus ou au-dessous de cette plage. Le niveau de crête vrai devrait être maintenu sous -2 dB TP afin de fournir la marge nécessaire pour éviter l'écrêtage potentiel à la suite d'un traitement aval (p. ex. le codage audio utilisé dans la distribution).

Point à retenir : Dans la production de contenu, lorsque le niveau sonore de la distribution du programme est inconnu ou n'a pas été spécifié, mixer le Niveau de dialogue à -24 LKFS avec crêtes vraies sous -2 dB TP.

I.8 EXIGENCE DE LA FCC

Le document ATSC A/53, partie 5:2010 [1] rend obligatoire le transport du paramètre dialnorm et de valeurs dialnorm réglées de façon appropriée.

Point à retenir : Le paramètre dialnorm de l'encodeur AC-3 de la station sera réglé pour qu'il corresponde à l'intensité sonore du Niveau de dialogue moyen du contenu.

¹³ Ou lorsqu'il n'existe pas d'accord préalable entre les parties pour ce qui a trait à l'intensité sonore.

I.9 MESURE DE CONTENU POSTPRODUIT

Voir la section 5.2.1 de la PR.

I.9.1 Contenu longue durée

Une partie représentative du contenu dominée par le dialogue type (c.-à-d. non par des cris ou des chuchotements) devrait être isolée et mesurée durant ou après le mixage final. En l'absence de dialogue, l'intensité sonore de l'élément du contenu sur lequel se fonderait un téléspectateur raisonnable pour régler le volume devrait être mesurée. Si aucune des deux techniques n'est possible ou pratique, l'intensité sonore de tout le contenu devrait être mesurée. Si le contenu contient des périodes importantes de silence, consulter la section 5.2.3 de la PR.

Point à retenir : Mesurer l'intensité sonore de passages du contenu longue durée où il y a des dialogues types et enregistrer la valeur comme Niveau de dialogue du contenu.

I.9.2 Contenu courte durée

Voir la section 5.2.4 de la PR.

Point à retenir : Mesurer l'intensité sonore de tous les canaux audio¹⁴ et de tous les éléments intégrés de la piste sonore, sur toute la durée du contenu de courte durée.

¹⁴ Le canal LFE n'est pas compris.

Annexe J :

Exigences relatives à l'établissement et au maintien de l'intensité sonore des messages publicitaires à la télévision numérique

J.1 INTRODUCTION ET PORTÉE

Les recommandations présentées dans cette annexe sont fondées sur d'autres sections de la Pratique recommandée. Cette annexe contient la marche à suivre nécessaire pour contrôler efficacement l'intensité sonore des messages publicitaires destinés à la télévision numérique.

J.2 GESTION DE L'INTENSITÉ SONORE

Le but de l'opérateur est de présenter aux téléspectateurs une intensité sonore uniforme entre les programmes, les messages publicitaires, le matériel de promotion et les messages d'intérêt public. Dans la mesure du possible, l'opérateur devrait offrir aux téléspectateurs le son le plus précis, de la plus haute qualité et libre de tout phénomène parasite ne faisant pas partie du produit original livré par le fournisseur de contenu.

J.3 PARAMÈTRE dialnorm POUR LES SYSTÈMES AC-3

Le document ATSC A/53, partie 5:2010 [1] section 5.5 décrit le transport du paramètre dialnorm et de valeurs dialnorm réglées de façon appropriée¹⁵.

J.4 MESURE DE L'INTENSITÉ SONORE DES MESSAGES PUBLICITAIRES

Lorsque l'intensité sonore d'un contenu courte durée (p. ex. message publicitaire) est mesurée, il est vital qu'elle soit mesurée en unités LKFS et que tous les canaux audio¹⁶ et tous les éléments de la piste sonore soient mesurés sur toute la durée du contenu.

J.5 MESSAGE PUBLICITAIRES AU POINT D'INSERTION

Dans le cas de l'insertion de contenu courte durée¹⁷ (p. ex. message publicitaire), il est vital que l'intensité sonore du contenu courte durée inséré, mesurée suivant les indications de la section J.4, corresponde¹⁵ au réglage dialnorm de ce flux audio AC-3 inséré, conformément à la section J.3.

¹⁵ Voir la section 7.1 du présent document.

¹⁶ Conformément à la recommandation BS.1770, le canal LFE n'est pas compris.

¹⁷ Voir la section 8.4 du présent document.

Annexe K :**Exigences relatives à l'établissement et au maintien de l'intensité sonore des messages publicitaires à la télévision numérique quand on utilise des codecs non-AC-3****K.1. INTRODUCTION ET PORTÉE**

Les recommandations présentées dans cette annexe sont fondées sur d'autres sections de la Pratique recommandée. Cette annexe contient la marche à suivre nécessaire pour contrôler efficacement l'intensité sonore des messages publicitaires destinés à la télévision numérique quand des codecs audio non-AC-3 sont utilisés.

K.2. GESTION DE L'INTENSITÉ SONORE

Le but de l'opérateur est de présenter aux téléspectateurs une intensité sonore uniforme entre les programmes, les messages publicitaires, le matériel de promotion et les messages d'intérêt public.

K.3. INTENSITÉ SONORE DES CANAUX NON-AC-3

Il est vital que la chaîne de distribution fonctionne suivant une valeur de intensité sonore cible sélectionnée par l'opérateur (et mesurée en unités LKFS) pour le contenu présenté sur la chaîne.

K.4. MESURE DE L'INTENSITÉ SONORE DES MESSAGES PUBLICITAIRES

Lorsque l'intensité sonore d'un contenu courte durée (p. ex. message publicitaire) est mesurée, il est vital qu'elle soit mesurée en unités LKFS et que tous les canaux audio¹⁸ et tous les éléments de la piste sonore soient mesurés sur toute la durée du contenu.

K.5. MESSAGE PUBLICITAIRE AU POINT D'INSERTION

Dans le cas de l'insertion de contenu courte durée (p. ex. message publicitaire), il est vital que l'intensité sonore du contenu courte durée inséré, mesurée suivant les indications de la section K.4, corresponde à la valeur de intensité sonore cible de la chaîne de distribution à ± 2 dB près.

¹⁸ Conformément à la recommandation BS.1770, le canal LFE n'est pas compris.



Pratique recommandée par l'ATSC : Techniques d'établissement et de maintien de l'intensité sonore pour la télévision numérique

Document A/85:2011, 25 juillet 2011

Advanced Television Systems Committee, Inc.
1776 K Street, N.W., Suite 200
Washington, D.C. 20006