

A1

**DEMANDE
DE BREVET D'INVENTION**

(21)

N° 81 24616

(54) Procédé perfectionné d'exploration sismique.

(51) Classification internationale (Int. Cl. 3). G 01 V 1/28, 1/38; H 03 K 13/22.

(22) Date de dépôt 31 décembre 1981.

(33) (32) (31) Priorité revendiquée : *EUA, 31 décembre 1980, n° 221,559.*

(41) Date de la mise à la disposition du
public de la demande B.O.P.I. — « Listes » n° 26 du 2-7-1982.

(71) Déposant : Société dite : MOBIL OIL CORPORATION, résidant aux EUA.

(72) Invention de : Aaron Jo Davis.

(73) Titulaire : *Idem* (71)

(74) Mandataire : Cabinet Beau de Loménie,
55, rue d'Amsterdam, 75008 Paris.

La présente invention concerne un procédé d'exploration sismique et, plus particulièrement, un procédé de compression de grandes quantités de données numériques résultant d'une exploration sismique, qui permet qu'on puisse les transmettre et, ou bien, les stocker plus efficacement.

C'est une pratique courante depuis de nombreuses années d'effectuer des explorations sismiques pour la recherche de pétrole, de gaz et d'autre minéraux, les techniques typiques consistant à produire une onde acoustique à la surface de la terre ou de l'océan. Cette onde se déplace vers le bas dans la terre et est renvoyée vers le haut par des couches de roche souterraines en direction de la surface de la terre ou de l'océan où son retour peut être détecté. De façon typique, le signal de sortie des détecteurs est constitué de signaux électriques analogiques qui sont mis sous la forme numérique et sont enregistrés. Les signaux enregistrés peuvent être ensuite analysés et utilisés pour produire une image de la structure souterraine de la terre, laquelle peut être interprétée par des géophysiciens en vue de la recherche de pétrole, de gaz ou d'autres minéraux intéressants.

Dans ce domaine de l'exploration sismique, il est souhaitable de produire des images toujours meilleures, ce qui nécessite l'enregistrement de données numériques plus nombreuses. A l'heure actuelle, un moyen typique d'exploration à partir d'un navire met en jeu l'enregistrement d'échantillons provenant de 208 géophones remorqués, chacun étant échantillonné toutes les 4 millisecondes avec précision de 16 bits. L'enregistrement de ces données s'effectue à un rythme tel que toute une bobine de bande magnétique se remplit de ces données en 10 minutes environ. Il est clair qu'il serait souhaitable de réduire la quantité des données à stocker, si ceci pouvait être réalisé sans perte de précision.

Une autre direction de développement des schémas de compression de données utiles se rapporte au fait qu'il est fréquemment souhaitable que ces données soient transmises du navire d'exploration océanique à un calculateur d'analyse se trouvant à la base, tandis que le navire océanique reste dans la zone

d'exploration. De cette manière, si les données sont particulièrement intéressantes, ou bien si, pour une raison quelconque, elles ont été enregistrées de manière inadéquate et nécessitent une nouvelle exploration, le navire peut revenir sur la zone d'exploration sans devoir effectuer un voyage d'une longueur excessive. Toutefois, transmettre des représentations numériques complètes des données sismiques enregistrées est extrêmement coûteux dans le contexte actuel des installations de transmission existantes, et cette situation n'est pas susceptible de s'améliorer. Voilà donc un autre domaine où la compression de données sismiques pourrait se révéler utile.

De façon claire, pour que des techniques de compression de données soient utiles, il ne faut pas qu'elles autorisent une déformation trop importante du signal au moment de la décompression. Deux facteurs particuliers sont mis en jeu; tout d'abord la différence quadratique moyenne entre la trace du mot complet et cette même trace après la compression et la décompression; et, en deuxième lieu, la corrélation de la trace de différence avec une trace à l'exactitude du mot entier. L'erreur quadratique moyenne décrit quantitativement la distorsion présente dans les traces comprimées et décomprimées. La corrélation indique si les erreurs faites s'accumulent (c'est-à-dire si les erreurs se réduisent avec l'emploi de moyens de la technique antérieure visant à réduire le bruit aléatoire dans les enregistrements sismiques par sommation) et, si ce n'est pas le cas, qu'elle est la forme de la distorsion.

Les moyens de la technique antérieure permettant la compression de données dans des systèmes sismiques n'ont pas réussi à produire des résultats convenables. Ces moyens techniques se classent suivant deux catégories de base. Une catégorie concerne les techniques du domaine de la fréquence, comme le montre le mieux possible l'article "Seismic Data Compression Methods" de L.C. Woods dans Geophysics, 30, n° 4 (1974) p. 499 à 525, et l'autre catégorie concerne les techniques du domaine du temps. Il a été fait usage de la quantification du seul bit de gain, c'est-à-dire de la caractérisation du mot complet par l'exposant de la puissance de 2 la plus proche de l'échantillon de façon qu'il faille transmettre seulement

l'exposant. En clair, ceci peut impliquer des erreurs égales à la moitié de la différence entre puissances de 2 successives. De plus, les erreurs sont corrélées à l'amplitude du signal, puisque, lorsque l'amplitude du signal augmente, cela est également vrai pour

5 l'erreur probable. Deux variantes de la méthode de modulation par impulsions codées différentielles (connue sous l'appellation méthode DPCM, que nous utiliserons dans cette description) ont également été essayées, à savoir la méthode DPCM simple et la méthode DPCM à somme courante. Ces techniques ont en général tenté d'utiliser

10 le fait que la différence entre échantillons successifs est plus petite que les échantillons eux-mêmes. Toutefois, la méthode DPCM à somme courante, bien qu'elle soit très rapide, entraîne un bruit quadratiquement moyen plutôt élevé. La méthode DPCM simple présente une distorsion plus réduite mais nécessite, pour l'élimination des

15 erreurs de propagation, un long filtre à coupure basse. C'est pourquoi, tandis que les techniques DPCM se révèlent utiles, il subsiste un besoin d'amélioration dans les techniques sismiques utilisées.

L'application des techniques DPCM et du codage de la parole est discutée de façon générale par N.S. Jayant dans

20 "Digital Coding of Speech Waveforms : PCM, DPCM and DM Quantizers", Proceedings of the IEEE, 62, n° 5 (mai 1964), mais en relation avec le codage de formes d'onde de parole et non avec le codage de données sismiques.

On notera que les données sismiques ont un caractère généralement répétitif, c'est-à-dire qu'elles impliquent une

25 onde sinusoïdale en convolution avec une série de réflectivité. On peut utiliser, pour la compression de données, la nature ondulatoire des enregistrements en tant que telle, et il n'est pas nécessaire de supposer que les données sont aléatoires. C'est un

30 fait qui est exploité par les techniques de modulation par impulsions codées différentielles mentionnées ci-dessus. Il existe trois sortes d'avantages aux techniques de modulation par impulsions codées différentielles par rapport au codage du seul bit de gain ou au codage du bit de signe, dans lequel un signe seulement est transmis,

35 indiquant si un échantillon donné est positif ou négatif. D'abord, les erreurs sont plusieurs fois plus petites que celles se produisant

lorsqu'on quantifie la forme d'onde complète en un nombre fixe de bits de gain binaires. En deuxième lieu, les erreurs sont beaucoup moins corrélées que pour le codage du bit de gain ou le codage du bit de signe, de sorte qu'il est moins probable que les erreurs
5 se renforcent au moment du traitement des données sismiques. En troisième lieu, les techniques de modulation par impulsions codées différentielles sont moins sensibles au bruit de basse fréquence que ce n'est le cas pour les techniques de codage du bit de gain ou du bit de signe et elles sont donc beaucoup plus appropriées
10 au filtrage f-k, qui est une technique bien connue pour réduire le bruit et que l'on emploie couramment dans le traitement des données sismiques. Ainsi, il est souhaitable d'utiliser une certaine forme de modulation par impulsions codées différentielles, mais, ainsi que cela a été mentionné ci-dessus, ni la technique DPCM simple
15 ni la technique DPCM à somme courante ne produisent de résultats à distorsion suffisamment faible pour que ces techniques soient utiles.

L'invention vise à proposer un procédé d'exploration sismique dans lequel des données peuvent être enregistrées sur place
20 et transmises à une base en vue du traitement, ce procédé étant économique du point de vue du temps de transmission mais ne sacrifiant pas la précision. De cette manière, les résultats du traitement peuvent être renvoyés sur le terrain pendant l'exploration de façon que, s'il est nécessaire de réexaminer la zone ou si une
25 zone se révèle être d'un intérêt particulier demandant une étude supplémentaire, la décision appropriée puisse être prise pendant que l'équipe d'exploration se trouve encore au voisinage du lieu intéressant, qu'il s'agisse d'une équipe terrestre ou d'une équipe embarquée à bord d'un navire. Selon l'invention, ceci peut être
30 réalisé au moyen d'un procédé perfectionné de codage en modulation par impulsions codées différentielles permettant une compression des données sismiques.

L'invention propose donc un procédé d'exploration sismique dans lequel l'énergie acoustique est transmise au travers
35 d'une structure souterraine, l'énergie réfléchie par une interface souterraine et renvoyée vers le haut jusqu'à la surface est mesurée,

des signaux de données de sortie électriques sont produits en réponse à l'énergie ainsi mesurée, et les signaux de données sont comprimés en vue du stockage et, ou bien, de la transmission, l'opération de compression de données consistant à :

- 5 échantillonner les signaux à un taux fixe;
 analyser plusieurs échantillons de sortie successivement en fonction d'un opérateur de prédiction pour produire une valeur prédite de l'échantillon d'entrée suivant;
 comparer l'échantillon d'entrée suivant avec la
- 10 valeur prédite correspondante afin de produire une valeur d'erreur de prédiction;
 quantifier la valeur d'erreur de prédiction de manière à produire un nombre de quantum; et
 memoriser et, ou bien, transmettre le nombre de
- 15 quantum;
 et l'opération de compression de données étant telle que, lors de l'inversion de l'opérateur en vue de la décompression des nombres de quantum stockés, une réaction puisse être appliquée à l'opérateur inversé pour faire que le bruit contenu dans les
- 20 échantillons quantifiés passe sans filtrage.
 Ainsi l'invention surmonte les inconvénients des systèmes connus décrits ci-dessus en utilisant un schéma de compression de données sismiques qui utilise un codage en modulation par impulsions codées différentielles à prédiction linéaire exploitant la nature de convolution des données sismiques. Dans ce procédé,
- 25 un nombre donné d'échantillons est quantifié et analysé afin de produire un échantillon suivant prédit. La différence entre l'échantillon suivant réel et l'échantillon suivant prédit est alors codée et transmise. L'échantillon quantifié devient alors le dernier des
- 30 échantillons précédents, et est utilisé pour produire l'échantillon réel suivant. La technique est désignée comme étant un codage de formes d'onde à prédiction linéaire. Elle est utilisée en relation avec un schéma de quantification souple dans lequel la différence entre les échantillons prédit et réel est comparée avec un nombre
- 35 prédéterminé de "puits" ou "fenêtres" dans lesquels on peut s'attendre à ce que la valeur du signal se trouve. Les fenêtres sont dilatées

ou contractées en fonction des résultats de la comparaison. De cette manière, le schéma de codage à prédiction linéaire avec quantification souple d'intervalles uniformes produit l'erreur quadratique moyenne la plus petite ainsi que la corrélation la plus aléatoire

5 des erreurs avec les données, par comparaison avec une technique de quantification qui s'adapte elle-même au signal et provoque ainsi une corrélation des erreurs avec le signal. Le schéma de prédiction linéaire permet également la conception d'un format de

10 présentation des données de longueur fixe. De plus, le procédé de codage de prédiction linéaire n'est pas sensible aux erreurs de transmission puisque l'erreur de prédiction est répartie sur la longueur de l'opérateur, c'est-à-dire sur les échantillons précédents comparés avec l'échantillon suivant. Le schéma de modulation

15 par impulsions codées différentielles à prédiction linéaire selon l'invention dépend de la largeur de bande du signal (c'est-à-dire que plus la largeur de bande est grande, plus l'erreur de prédiction est grande, et, ainsi, plus l'erreur de quantification est grande). Le fait de faire des coupes parmi les données, par exemple en prenant un échantillon sur deux, réduit le volume des données d'un

20 facteur 2, mais augmente également l'erreur de quantification. Heureusement, les données sismiques ont une composante périodique passe-bas relativement invariable et peuvent donc être comprimées au moyen d'une prédiction linéaire, puisque c'est la nature sinusoïdale de l'onde sismique qui permet que la valeur prédite soit

25 relativement proche de la valeur réelle, en autorisant qu'un signal de différence relativement petit soit codé et transmis avec une erreur minimale.

Alors que, dans de nombreux cas, les signaux d'erreur codés sont eux-mêmes intéressants sans décodage ni décompression, à la réception du signal codé transmis, l'opération

30 inverse peut donc être réalisée pour produire une forme d'onde de sortie sensiblement identique à celle qui a été codée. Une réaction peut être utilisée pendant la décompression pour filtrer le bruit, ce qui améliore le rapport signal-bruit.

35 La description suivante, conçue à titre d'illustration de l'invention, vise à donner une meilleure compréhension de

ses caractéristiques et avantages; elle s'appuie sur les dessins annexés, parmi lesquels :

- la figure 1 représente une vue générale d'un système d'exploration sismique selon l'invention;
- 5 - la figure 2 montre une comparaison graphique entre le codage par mot complet de signaux sismiques et quatre procédés de codage de compression considérés;
- la figure 3 représente les différences entre le codage par mot complet et les quatre procédés de codage de compression présentés sur la figure 2;
- 10 - la figure 4 représente les cinq enregistrements de la figure 2 après le filtrage f-k et le traitement de déconvolution de pointe ;
- la figure 5 représente les différences existant
- 15 entre le codage par mot complet et les quatre autres techniques de codage considérées après le filtrage f-k et le traitement de déconvolution de pointe ;
- la figure 6 représente les enregistrements de la figure 4 après correction dynamique et sommation;
- 20 - la figure 7 montre les différences existant entre l'échantillon sismique à mot complet ayant subi la correction dynamique et la sommation et les quatre enregistrements codés;
- la figure 8 présente un organigramme simplifié d'un processus de quantification souple utilisé dans le procédé
- 25 de l'invention;
- la figure 9 montre schématiquement comment le procédé de quantification souple de la figure 8 fonctionne;
- la figure 10 montre le caractère inversible de la déconvolution utilisée dans la technique de l'invention;
- 30 - la figure 11 montre la manière dont l'opérateur de codage à prédiction linéaire peut être utilisé pour réduire l'erreur;
- la figure 12 montre comment une réaction peut être appliquée au système général présenté sur la figure 10;
- 35 - la figure 13 montre un organigramme résumé de la technique de compression utilisée selon l'invention;

- la figure 14 montre un organigramme résumé de la technique de décompression utilisée selon l'invention;
 - la figure 15 présente une visualisation de données sismiques réelles utilisant un codage par mot complet;
 - 5 - la figure 16 montre les mêmes données qui ont été codées pour deux bits par mot et ont été décompressées au moyen du procédé de l'invention;
 - la figure 17 présente un organigramme détaillé du programme de formation de l'opérateur;
 - 10 - la figure 18 montre un organigramme détaillé du programme de codage;
 - la figure 19 montre un grossissement de la partie quantification du programme de codage présenté sur la figure 18; et
 - les figures 20a et 20b montrent des organigrammes
 - 15 détaillés décrivant la décondensation et le décodage effectués sur un train de données comprimé après transmission ou stockage.
- On se reporte maintenant aux dessins et, en particulier, à la figure 1, qui présente, de façon générale, un système d'exploration sismique sous-marin selon l'invention (bien qu'il ne
- 20 faille pas oublier que l'invention s'applique tout aussi bien à des opérations terrestres). Un navire d'exploration 10 porte plusieurs sources 12 d'énergie acoustique, comme des canons à air comprimé, qui sont destinées à émettre une onde acoustique 14 vers le bas dans la mer. L'onde passe au travers du fond de l'océan
- 25 et de la première couche sous-marine 16 pour être finalement réfléchie vers le haut par une couche 18 située plus bas et atteindre plusieurs hydrophones 20 remorqués derrière le navire. Les signaux de sortie des hydrophones 20 sont transmis par un câble 22 à un équipement de traitement classique monté à bord du navire d'exploration 10, et constitué par exemple d'amplificateurs 24 et de
- 30 filtres 26. Les données peuvent être converties à la forme numérique dans un convertisseur analogique-numérique 28, puis peuvent être codées selon l'invention dans un processeur de codage 30. Les données comprimées peuvent être stockées et, ou bien, transmises
- 35 par l'intermédiaire d'un émetteur 32 et d'une antenne 34, éventuellement par l'intermédiaire d'un satellite 36, à destination d'une

antenne 38 terrestre, en vue d'être peut être encore traitées dans un ordinateur indiqué de façon générale par la référence 40.

Ainsi que cela a été noté ci-dessus, des quantités importantes de données sont produites pendant les explorations sismiques terrestres aussi bien que marines, et il est donc souhaitable de disposer de procédés de compression de données de sorte que la quantité de données à transmettre soit limitée. Comme cela a également été noté ci-dessus, les données sismiques ont un caractère généralement ondulatoire de sorte que les échantillons successifs de signaux de sortie analogiques des hydrophones 20 ne varient pas fortement de l'un au suivant. Ainsi, par exemple, un schéma de modulation par impulsions codées différentielles simple peut être utilisé pour coder la différence existant entre deux échantillons successifs, au lieu de coder toute l'amplitude de l'échantillon de données. On considère typiquement que cette différence est inférieure à l'amplitude globale et que des bits sont ainsi économisés. Lorsque la corrélation entre deux échantillons est élevée (supérieure à 0,5 environ), les différences entre échantillons successifs sont petites.

Alors que ce procédé peut valablement être mis en oeuvre, il n'assure pas l'absence d'erreurs résiduelles de basse fréquence qui tendent à être fortement amplifiées au moment de la reconstruction par intégration dans le poste de traitement terrestre 40. Ces basses fréquences amplifiées provoquent un glissement et quelquefois un déplacement en courant continu qui peuvent se révéler instables lors des opérations de traitement ultérieures. Il est ainsi nécessaire d'appliquer avant l'intégration un long filtre à coupure basse aux échantillons et de retirer la moyenne ensuite. Malheureusement, ceci produit une distorsion indésirable dans le premier étage de traitement, celle-ci étant en outre aggravée par les étages ultérieurs.

Un perfectionnement à la modulation simple par impulsions codées différentielles numériques, connu sous l'appellation de modulation d'impulsions codées différentielles à somme courante, fait appel à la récurrence pour réduire le problème de basse fréquence mentionné ci-dessus. Le but de ce système est

d'ajuster de façon continue les basses fréquences en prenant la différence entre le signal entrant et la somme courante sur toutes les différences quantifiées précédentes. S'il n'y a pas d'erreur dans la quantification, cette opération se ramène simplement à

5 prendre la trace de différence, comme ci-dessus. Toutefois, puisque des erreurs peuvent être attendues dans le processus de quantification, la différence à quantifier est le changement apparaissant entre deux signaux d'entrée successifs plus l'erreur de quantification de la différence précédente. A chaque fois qu'une importante

10 erreur se produit, elle est donc corrigée par la différence suivante au lieu de se propager dans toute la trace. Alors que cette technique produit un meilleur effet que le fait de simplement faire la différence des données d'un point de vue de propagation d'erreur et permet d'éliminer le long filtre à coupure basse et l'enlèvement de

15 la moyenne mentionnés ci-dessus, il existe une corrélation de l'erreur avec les données.

Par souci de clareté, le procédé qui est ordinairement appelé procédé de modulation par impulsions codées différentielles est ici appelé procédé de modulation par impulsions codées différentielles à prédiction linéaire. Par comparaison avec la modulation

20 par impulsions codées différentielles à somme courante, au lieu de la prédiction du dernier échantillon au moyen de la somme des échantillons précédents, il est tenté, selon ce procédé, de prédire l'échantillon suivant à partir d'une combinaison linéaire, commandée par un opérateur de prédiction, de plusieurs échantillons

25 passés. Par exemple, si la quantification est exempte d'erreur, la différence entre le signal d'entrée courant et l'échantillon de prédiction est simplement l'erreur de prédiction. Une telle séquence est particulièrement souhaitable dans un but de quantification, puisque le spectre est "blanchi" (c'est-à-dire a une largeur de bande plus grande) et que le niveau d'amplitude quadratique moyen est minimisé. Toutefois, la quantification introduit un bruit dans le système, lequel doit être maîtrisé au moment de la reconstruction pour ne pas rendre celle-ci non fiable. Selon le procédé de l'in-

30 vention, on maîtrise le bruit de quantification en appliquant une réaction du signal de sortie reconstruit sur le signal d'entrée

précédent afin de corriger l'opérateur de prédiction au lieu de simplement prédire sur la base d'une série de signaux d'entrée.

La figure 2 montre la comparaison de données d'entrée avec des données de sortie produites au moyen de quatre des techniques discutées ci-dessus. Les cinq enregistrements sont comparables à des sismogrammes, c'est-à-dire que leur axe vertical est le temps tandis que l'axe horizontal est la distance comptée depuis une origine donnée. La partie a de la figure 2 représente la visualisation de données synthétiques codées suivant l'échantillon de mot de 32 bits complet; il s'agit du type de données qui doit être comprimé selon le procédé de l'invention. La partie b de la figure 2 montre les données codées de mot complet de la partie a après leur codage au moyen de techniques de modulation par impulsions codées différentielles à prédiction linéaire. La partie c de la figure 2 est comparable et utilise la modulation par impulsions codées différentielles à somme courante. La partie d de la figure 2 montre la modulation par impulsions codées différentielles simple. Enfin, la partie e de la figure 2 montre le codage du seul bit de gain. L'homme de l'art admettra que les cinq sismogrammes présentés sur la figure 2 sont sensiblement identiques dans leur tracé, mais que la qualité du codage diminue notablement de la partie b à la partie e de la figure 2. Ceci est représenté graphiquement sur la figure 3, qui représente les résultats respectivement obtenus lorsqu'on soustrait des données de mot complet de la partie a de la figure 2 les données codées présentées sur les parties b à e de la figure 2, ceci fournissant une indication de la précision du schéma de codage employé. La partie a de la figure 3, correspondant à la modulation par impulsions codées différentielles à prédiction linéaire est nettement la meilleure des techniques de codage utilisées; l'erreur significative se limite à une région comprise entre 1,8 et 2,3 s. On peut voir que c'est seulement lorsque les amplitudes sont élevées et que le caractère de prédiction est faible, par exemple pour la superposition de tous les événements sur les traces supplémentaires, que les erreurs significatives de la représentation en modulation par impulsions codées différentielles à prédiction linéaire.

La figure 4 montre des sismogrammes analogues à ceux présentés sur la figure 2, mais après que des opérations de

traitement de données classiques ont été effectuées, à savoir des méthodes de déconvolution de pointe et de filtrage f-k. Ces deux méthodes s'appuient sur l'utilisation de processus mathématiques permettant de produire une information de fréquence cohérente à 5 partir d'enregistrements sismiques. Dans ce cas, l'"explosion aérienne", c'est-à-dire le bruit de surface produit par l'énergie acoustique, a d'abord été éliminé par filtrage f-k; une déconvolution de pointe a ensuite été appliquée aux enregistrements d'entrée. La 10 partie a de la figure 4 montre qu'il y a très peu de bruit de haute fréquence et un petit effet d'explosion aérienne rémanent qui subsistent dans le système de codage de mot complet. De même, avec les trois procédés de modulation par impulsions codées différentielles, respectivement présentés sur les parties b, c et d de la figure 4, l'effet d'explosion aérienne est atténué de manière satisfaisante. 15 Toutefois, le codage du seul bit de gain comporte, en plus d'un niveau de bruit supérieur, un effet d'explosion aérienne rémanent important. Le procédé de modulation par impulsions codées différentielles simple (partie d de la figure 4) a été amélioré par rapport au procédé à somme courante de la partie c de la figure 4 par une 20 opération de déconvolution qui supprime certains effets du filtre à coupure basse indiqués ci-dessus.

Les enregistrements de différence de la figure 5 sont comparables à ceux de la figure 3, c'est-à-dire que les parties b, c, d et e de la figure 4 ont été respectivement soustraites de la 25 partie a de la figure 4 pour produire les parties a, b, c et d de la figure 5. De nouveau, le procédé de modulation par impulsions codées différentielles à prédiction linéaire correspondant à la partie a de la figure 5 présente l'erreur la plus petite; toutefois, il apparaît que l'essentiel de l'erreur se trouve tout au long des 30 événements de réflexion qui ont subi la déconvolution. La modulation par impulsions codées différentielles simple présente également un signal relativement bon (partie c de la figure 5) et, encore une fois, le codage du seul bit de gain (partie d de la figure 5) comporte un bruit important. La comparaison des rapports signal-bruit 35 de ces enregistrements supporte une telle conclusion. Le rapport signal-bruit quadratique moyen des signaux en modulation par

impulsions codées différentielles à prédiction linéaire pris en moyenne sur 25 traces est de 3,93; celui du procédé à somme courante est de 1,43; celui de la modulation simple par impulsions codées différentielles est de 2,03 et celui du codage du bit de gain seul est de 1,05.

Un autre moyen bien connu de la technique antérieure permettant d'améliorer le rapport signal-bruit dans des données sismiques est obtenu par la correction dynamique et la sommation. Avec ces processus, les enregistrements subissent une correction temporelle pour compenser la variation des déports horizontaux entre la source et le récepteur et subissent une sommation, c'est-à-dire sont sommés mathématiquement ou électriquement, de sorte que le bruit, qui est aléatoire, tend à s'annuler, tandis que les signaux sont amplifiés. La figure 6 présente, de gauche à droite, des exemples de cette technique appliquée aux cinq enregistrements d'essai des parties a à e de la figure 4, tandis que la figure 7 montre des traces de différence des cinq traces de la figure 6, c'est-à-dire où les quatre traces codées ont été soustraites de la trace de mot complet de la figure 6. Après application de cette technique, le rapport signal-bruit de quantification des signaux traités par le procédé de codage par impulsions différentielles à prédiction linéaire est de 5,7; celui à somme courante est de 2,63; celui de la modulation par impulsions codées différentielles simple est de 3,17; et celui du codage du bit de gain seul est de 1,47. Puisqu'aucun de ces rapports n'a été amélioré par la correction dynamique et la sommation d'un facteur supérieur à 1,8 (pour le cas de la modulation par impulsions codées différentielles à somme courante), on peut conclure que les erreurs de toutes ces techniques ne sont pas sans corrélation avec les données. Alors que la modulation par impulsions codées différentielles à prédiction linéaire présente la distorsion la plus petite après traitement, on peut déduire des traces de différence de la figure 7 que les erreurs apparaissent être moins corrélées avec les données après sommation. Même si la modulation simple par impulsions codées différentielles présente un rapport signal-bruit plus élevé que la modulation par impulsions codées différentielles à somme courante, la distorsion introduite par le filtre à coupure basse est beaucoup plus remarquable sur les traces sommées de la méthode

DPCM simple que sur celle traitée selon la technique à somme courante. La trace de différence relative au codage du bit de gain seul laisse voir tous les événements majeurs, mais le niveau de bruit est très élevé et chaque ondulation présente une distorsion notable.

- 5 Ainsi que cela est discuté ci-dessus, il peut être montré que le codage par modulation par impulsions codées différentielles à prédiction linéaire peut être, si l'on suppose un canal sans bruit, appliqué avec succès à des données sismiques à distorsion minimale. Un même résultat peut être démontré pour des
- 10 données réelles. De plus, on démontre que les schémas moins complexes de modulation par impulsions codées différentielles simple et à somme courante produisent de meilleurs résultats que le codage du seul bit de gain, mais se trouvent encore au-dessous de la qualité obtenue avec la modulation par impulsions codées différentielles
- 15 à prédiction linéaire. Un aspect important de l'invention se rapporte à un perfectionnement supplémentaire apporté à la modulation par impulsions codées différentielles à prédiction linéaire. Il est utilisé un schéma de codage à longueur de mot fixe qui, au contraire du codage du gain seulement, implique un taux de compression spécifique. Puisque les données sismiques demandent une possibilité de
- 20 fixation automatique des limites du gain, le schéma de codage souple s'impose. Selon ce schéma de codage souple, les fenêtres de quantification dans lesquelles toutes les données doivent s'ajuster varient en fonction de l'amplitude globale du signal. En termes
- 25 simples, lorsque le schéma de codage souple est appliqué à la modulation par impulsions codées différentielles à prédiction linéaire, ce schéma fait varier l'écartement des puits ou limites de valeurs possibles dans lesquels la trace d'erreur de prédiction tombe et dilate ou bien contracte ces limites en fonction du signal de
- 30 sortie précédent. Par exemple, si le dernier signal de sortie indique que l'erreur se trouve dans le puits le plus à l'extérieur, le puits peut être dilaté pour le signal d'entrée suivant. Si l'échantillon suivant tombe dans un puits intérieur, le puits doit être contracté. De cette manière, on peut obtenir une meilleure adaptation de la
- 35 taille des puits avec l'erreur, tout en utilisant un nombre limité de puits; comme seule l'identification du puits doit être transmise et, ou bien stockée, on réalise une compression notable des données.

Ainsi, alors que le schéma de quantification souple est très économe vis-à-vis du nombre de bits nécessaire, il a encore la possibilité de s'ajuster à l'amplitude du signal. Il est particulièrement utile dans le cas d'un enregistrement de données sismiques, où l'amplitude diminue lentement en fonction du temps de même qu'elle varie graduellement d'un échantillon au suivant.

Le procédé de l'invention est illustré sous forme de blocs et de manière imagée sur les figures 8 et 9 respectivement. Initialement, il est fixé un nombre fini de "puits" qui tient compte de l'étendue approchée des valeurs attendues pour les quelques premiers échantillons de données. Le nombre de puits est égal à deux élevé à une puissance égale au nombre de chiffres binaires par mot spécifié. En d'autres termes, deux bits ($b=2$) permettent quatre puits ou états; trois ou quatre bits à transmettre par échantillon permettraient respectivement huit ou seize puits. Au fur et à mesure que l'échantillon $S(t)$ survient, l'identification $M(t)$ du puits qui loge l'échantillon est enregistrée. Ainsi, selon une formule ne dépendant pas du temps qui ne fait intervenir que le numéro du puits des échantillons précédents, la dimension du puits varie relativement au puits à l'intérieur duquel l'échantillon a été choisi. Ainsi, par exemple, si le signal tombe dans le plus grand puits, les puits sont dilatés pour l'échantillon suivant. Les dimensions des puits représentent des intervalles uniformes de sorte que tous les puits sont modifiés par un même facteur. Ces facteurs peuvent être déterminés pour des signaux à répartition normale avec divers coefficients de corrélation, ainsi que cela est établi dans l'article de N.S. Jayant indiqué ci-dessus, ainsi que dans son article "Adaptive Quantization with a One-Word Memory", Bell System Technical Journal, 52, n° 7, (septembre 1973) p. 1119 à 1144, ainsi que dans "Quantizing for Minimum Distortion" de J. Max dans IRE Transactions on Information Theory, 6, p. 7 à 12 (1960).

L'erreur E de tout échantillon particulier peut être écrite sous la forme :

$$E = |e| - |\hat{e}|,$$

où e est l'échantillon réel et \hat{e} est l'échantillon quantifié.

On choisit des valeurs absolues pour la définition, car le signe du signal est de préférence traité séparément, c'est-à-dire que l'on s'arrange pour que e et \hat{e} aient le même signe si e n'est pas égal à zéro. La quantification s'effectue de manière que tout e se trouvant dans l'intervalle

$$(n - 1) Q \leq e \leq nQ$$

soit quantifié sous la forme :

$$(n - 1/2) Q,$$

où Q est l'intervalle de quantification variable et n est le numéro de puits (1, 2, 3 ..., 2^{b-1} , 2^b). Lorsque N est égal à 2^b , le puits le plus grand, tout e se trouvant dans l'intervalle

$$(N - 1) Q \leq e \leq \infty$$

est quantifié sous la forme :

$$e = (N - 1/2) Q.$$

En supposant que l'échantillon réel ait une même probabilité de se trouver n'importe où dans l'intervalle, il s'ensuit que l'erreur E est approximativement proportionnelle à Q , l'intervalle de quantification. Puisque le schéma de quantification souple permet à Q d'augmenter et de diminuer lorsque le signal le fait, les erreurs (en supposant une distribution de probabilité uniforme dans l'intervalle) augmentent et diminuent également. Cette corrélation de l'amplitude des erreurs avec le signal entraîne l'apparition d'erreurs dans la largeur de bande spectrale du signal. Cette situation, qui est une conséquence de tout schéma de codage souple, est contraire au but visé selon lequel les erreurs doivent être petites et extérieures à la largeur de bande du signal, ainsi que cela a été discuté ci-dessus. La solution selon l'invention est de filtrer le signal avant la quantification de façon que le signal de sortie ait une bande aussi large que possible, avec une amplitude moyenne la plus faible possible. De cette manière, les variations de la taille des puits peuvent être minimisées, ainsi donc que les erreurs.

En plus des deux critères énoncés ci-dessus relatifs à l'opération de filtrage, c'est-à-dire la réduction de la variation d'amplitude et l'élargissement des spectres, il est nécessaire que tout filtrage soit inversible au moins dans le cas

5 ne comportant pas de bruit, de sorte que les données initiales puissent être retrouvées à partir de la séquence filtrée et codée. Il a été observé que le processus le plus sensible dans des opérations normales de traitement sismique est la déconvolution, c'est-à-dire la séparation d'événements répétitifs dans les données sismiques. Il s'ensuit que, si les enregistrements sismiques doivent

10 être altérés à une étape quelconque par codage, il sera préférable que se soit après la déconvolution. Le fait que la déconvolution puisse être inversée est un autre élément de jugement favorable. Le processus de déconvolution peut être décomposé de manière à être

15 inversible, ainsi que cela est présenté sur la figure 10. Les échantillons entrant $S(t)$ sont décomposés par sommation sur tous les n du produit $\hat{S}(t-n) \cdot A(n)$, où $A(n)$ est l'opérateur de prédiction. La comparaison de $S(t)$ et $\hat{S}(t)$ produit $R(t)$, l'écart de l'échantillon produit vis-à-vis de l'échantillon réel ou de l'"erreur prédite".

20 Un instant plus tard, l'opération inverse peut être effectuée sur $R(t)$ afin de produire le signal d'entrée exact $S(t)$.

Il est bien connu que, après déconvolution, la trace a un spectre plus large. C'est ce que l'on peut obtenir en trouvant l'opérateur de prédiction $A(n)$ qui, dans un mode de réalisation

25 préféré, est choisi de façon à minimiser la somme des carrés de la trace de l'erreur de prédiction. Un processus de minimisation pour cette variance est résumé sur la figure 11. La notation I désigne la somme de tous les carrés de tous les $R(t)$ dont chacun vaut $S(t) - \hat{S}(t)$, où $\hat{S}(t)$ est lui-même égal à la somme de tous les termes

30 $[S(t - n) \cdot A(n)]$. Ainsi, on peut résoudre une série de n équations linéaires pour $A(n)$ où

$$\frac{\partial I}{\partial A(n)} = 0,$$

afin de minimiser I . Dans un mode de réalisation particulièrement préféré de l'invention, la résolution du groupe résultant d'équations linéaires revient à résoudre une matrice de Toeplitz pour

35

le vecteur d'autocorrélation déplacé d'un échantillon en avant.

- Indépendamment de l'opération de préfiltrage qui est effectuée sur les données avant quantification, on note que, lorsque le filtre inverse est appliqué aux données reconstruites,
- 5 il affecte le bruit de quantification de la même manière que le signal. L'effet du filtre inverse est plus important sur des données quantifiées au moyen du schéma de codage DPCM simple. Même si les erreurs sont à peine remarquables dans la trace de différence, elles sont fortement amplifiées lors du filtrage inverse par l'intégration.
- 10 En conséquence, il faut employer un mécanisme de réaction pour commander le bruit. La nécessité de la réaction appelle, en retour, deux conditions supplémentaires sur l'opération de préfiltrage : ainsi, son opérateur doit avoir un retard minimal et être obtenu à partir d'une fenêtre de données à moyenne nulle. On connaît de
- 15 nombreuses méthodes pour obtenir un tel opérateur de prédiction, mais toutes impliquent un compromis entre le degré auquel l'erreur quadratique moyenne est minimisée et le taux de convergence de l'opérateur réciproque. Toutefois, il a été découvert que la méthode d'autocorrélation produit l'opérateur le plus efficace du point de
- 20 vue du calcul tout en conduisant à un opérateur à charge de face convenable, ce qui permet une régularisation de l'erreur de prédiction résultant du quantificateur souple. Puisque la quantification non linéaire amène un bruit de basse fréquence comme de haute fréquence, la méthode d'autocorrélation doit être appliquée à
- 25 l'autocovariance (moyenne retirée des données) plutôt qu'à l'autocorrélation en ce qui concerne la création de l'opérateur de prédiction de sorte que des décalages de basse fréquence ne soient pas prédits lors du traitement de données réelles. Ceci révèle un
- 30 contraste intéressant avec les techniques de compression de parole mentionnées ci-dessus. Dans la compression de la parole, les basses fréquences (c'est-à-dire inférieures à 200 Hz environ) peuvent être filtrées car elles contribuent peu au signal, et elles peuvent être écartées lors de la création de l'opérateur. Dans la compression de données sismiques, la gamme de fréquence typique est de
- 35 5 à 50 Hz si bien qu'une filtration à coupure basse élimine toutes les données intéressantes. Par conséquent, une technique d'autocorrélation telle que mentionnée ci-dessus doit être utilisée

(voir Rubiner et Schafer, "Digital Processing of Speech Signals", Prentice Hall, Englewood Cliffs, New Jersey (1978), en particulier le chapitre 8, section 8.1.1). En soustrayant la valeur moyenne de chacun des échantillons dans la fenêtre intéressante, on peut centrer
 5 sur zéro l'échantillon global, ce qui évite la prédiction d'un glissement de basse fréquence et un codage instable.

La figure 12 montre un organigramme de la boucle de réaction associée. En plus de la boucle présentée sur la figure 10, il est prévu une deuxième boucle de sommation qui comprend un noeud
 10 de sommation et une étape de prédiction produisant la formulation d'une réaction en $\hat{S}(t)$ pour le noeud de sommation principal. De cette manière, c'est un $\hat{R}(t)$ qui est transmis, c'est-à-dire qu'une erreur estimée, qui est calculée selon un processus définissable, pouvant être décodé du côté récepteur, ainsi que cela est également
 15 indiqué sur la figure 12, est transmise.

En résumé, le procédé de l'invention implique la résolution de plusieurs problèmes. Le premier problème est de minimiser le nombre d'états ou puits possibles dans lesquels les données peuvent être transférées. La solution consiste à utiliser un schéma
 20 de quantification souple; puisque le numéro du puits dans lequel un échantillon donné tombe est celui qui est transmis, il est tiré un avantage maximal du procédé de compression de données selon l'invention. De préférence, il est fait appel à un procédé de quantification souple à intervalle uniforme; ceci implique l'emploi d'une
 25 mémoire d'un seul mot retenant la dimension du puits d'un intervalle d'échantillonnage au suivant. Toutefois, la quantification souple produit un bruit qui est proportionnel au signal et qui a des composantes spectrales analogues. La solution à ce problème consiste à transformer le signal en une trace d'erreur de prédiction $R(t)$
 30 avant la quantification; toutefois, ceci soulève le problème que la transformée inverse du schéma d'erreur de prédiction amplifie le bruit et rétrécit le spectre. La solution à ce problème consiste à incorporer dans le système une boucle de réaction qui fait passer le bruit sans filtrage. L'organigramme résultant de l'application
 35 de ces concepts est celui présenté sur la figure 12.

Les figures 13 et 14 montrent respectivement les programmes de compression et de décompression. La compression des données, qui est représentée sur la figure 13, s'effectue au lieu de délivrance, ou à l'instant de stockage des données, et le programme de décompression correspondant est effectué à l'extrémité de réception, ou au moment de la récupération des données. L'organigramme de compression présenté sur la figure 13 montre que l'on calcule les opérateurs de prédiction, que l'on fixe les niveaux de quantification initiaux et que l'on calcule l'erreur, que l'on code celle-ci et que l'on condense la trace comprimée. Comme on peut le voir sur la figure 14, l'information A de l'opérateur, définissant les conditions initiales, devient un entête et est stockée ou transmise avec les codes de différence condensés $M(t)$, et le programme de décompression travaille sur eux. A cet égard, on note de nouveau que les données sismiques diffèrent notablement de la parole par exemple, en ce qu'elles ont une nature de convolution croissante; ceci permet de définir un seul opérateur $A(n)$ pour la totalité de la trace, de l'enregistrement, de la ligne ou de la région d'exploration. Par comparaison, dans le codage de la parole, l'opérateur doit être recalculé environ toutes les 20 millisecondes; en effet, l'opérateur représente typiquement tout ce qui est transmis et les erreurs de prédiction sont créées synthétiquement. Après le stockage et, ou bien la transmission, l'échantillon est décondensé de façon à produire $R(t)$, la donnée qui est décodée de façon à produire $R(t)$, et celle-ci est ajoutée à l'échantillon prédit qui est calculé en fonction des échantillons précédents de manière à produire l'échantillon reconstitué $\hat{S}(t)$. $\hat{S}(t)$ est utilisé pour prédire le signal de sortie suivant. L'inverse du schéma de compression présenté de manière imagée sur la figure 9 est donc utilisé pour réaliser l'inverse du changement de la taille des puits afin de préparer l'échantillon suivant $M(t)$. $R(t)$ est alors utilisé pour corriger le signal de sortie prédit, qui est donc l'échantillon reconstitué $\hat{S}(t)$.

L'homme de l'art comprendra donc qu'une formation correcte de l'opérateur de prédiction est importante pour la précision et la stabilité du signal obtenu par le procédé de l'invention.

Selon une particularité préférée de l'invention, les données utilisées pour calculer la fonction d'autocorrélation obéissent à une courbe fictive ou sont fixées pour un instant de départ constant. La longueur de la fenêtre employée vaut au moins 10 fois la longueur maximale de l'opérateur; si un enregistrement court se présente, la longueur de l'opérateur se réduit. La valeur moyenne est retirée de la fenêtre de données, mais non des données elles-mêmes. On calcule l'autocorrélation en supposant que tous les zéros se trouvent à l'extérieur de la fenêtre considérée. Le niveau initial du bruit blanc s'ajoute à l'ordonnée du centre afin d'éviter les difficultés résultant de l'inversion de zéros; pour calculer l'opérateur on utilise donc le programme de récurrence de Lévinson. La condition d'erreur est vérifiée et, si elle donne un résultat positif, il est calculé un autre opérateur auquel est ajouté 0,5% de bruit blanc en plus. Si la quantité de bruit blanc nécessaire est supérieure à 10%, la trace est "tuée", c'est-à-dire qu'elle n'est pas transmise ou enregistrée, selon le cas.

Le programme de compression peut lui-même être résumé de la manière suivante. La trace est d'abord initialisée par mise à zéro d'un premier nombre prédéterminé de points. Ainsi, la taille initiale des puits est fixée à une valeur égale à la moyenne d'un certain nombre prédéterminé de valeurs absolues non nulles rencontrées. Les facteurs de dilatation et de contraction sont fixés d'après l'article publié par Jayant en 1973 cité ci-dessus. L'erreur de prédiction est ensuite calculée, l'échantillon comprimé est codé et condensé, l'échantillon suivant est prédit, et ainsi de suite, de façon à répéter le calcul de l'erreur, le codage et la condensation pour chaque échantillon de données. La stabilité peut être confirmée de manière continue. Ainsi, le produit sortant du programme de compression de la figure 13 est une trace entière condensée contenant les composants aléatoires codés qui indiquent l'écart par rapport à l'erreur et l'information d'entête comportant l'opérateur et la taille initiale du puits.

Le programme de décompression est présenté sur la figure 14 et suit sensiblement le processus inverse. D'abord, les données sont décondensées pour être placées en représentation entière. Cette trace codée peut en fait être utilisée comme trace

de déconvolution égalisée du point de vue du gain ou bien les entiers peuvent être décodés pour produire une approximation de l'intervalle dynamique complet de la trace de déconvolution; enfin, naturellement, par l'utilisation de l'opérateur fourni dans l'entête, on peut reconstruire le signal d'entrée initial avec une très haute fidélité. L'homme de l'art admettra que le programme de compression selon l'invention, tout en produisant une réduction de 10 à 1 du temps nécessaire par exemple pour transmettre les données via une liaison radio par satellite ou pour mémoriser les données sur une bande magnétique, demande un certain temps de calcul pour sa mise en oeuvre. A cette fin, il est envisagé selon l'invention qu'un unique ordinateur soit de manière souhaitable réservé à ce but, c'est-à-dire soit programmé pour réaliser le programme de compression, par exemple à bord du navire. Il semble que l'emploi d'un minicalculetur tel que "PHOENIX I" comme unité de codage 30 (voir figure 1) soit acceptable.

Les figures 15 et 16 démontrent avec éclat le succès du procédé. La figure 15 est une représentation de données sismiques qui ont été stockées dans un ensemble de mots de 32 bits complets; la figure 16 quant à elle représente les mêmes données qui ont été comprimées à un niveau de 2 bits par échantillon, puis décomprimées de la manière indiquée ci-dessus. Les données de la figure 16 peuvent donc être transmises à un coût inférieur au dixième de celui de la figure 15. Toutefois, l'homme de l'art admettra que la figure 16 est très comparable à la figure 15 et se révèle tout à fait utile pour la détermination du succès d'opérations de prospection sismique en cours. De plus, il a été découvert que les variations du produit obtenu relativement au produit initial sont liées à l'échelle suivant laquelle les données sont graphiquement transcrites. Ce que montre la figure 16 correspond à douze traces par pouce (soit 2,5 cm environ) pour une vitesse de 5 pouces (soit 10 cm environ) par seconde. Pour une résolution inférieure, par exemple donnant plus de 20 traces par pouce, on ne note aucune différence entre l'enregistrement de mot complet et l'enregistrement comprimé-décomprimé de 2 bits. Pour l'échelle de douze traces par pouce portée sur les figures 15 et 16, on voit que la figure 16 présente effectivement une très légère

instabilité de haute fréquence intéressant les données sismiques de très grande longueur d'onde, ceci pouvant être facilement réduit par application d'un filtre à coupure haute aux données profondes de la coupe.

5 L'homme de l'art admettra également que, tandis que la décompression et la reconstitution des données dans leur format de présentation initial sont possibles, ceci n'est pas nécessaire avant qu'il soit fait usage des données. Les données sismiques peuvent donc être traitées comme entiers de 2 bits directement à partir
10 du format de présentation comprimé. Dans ce cas, l'égalisation de gain et la déconvolution avant sommation sont redondants dans le traitement sismique normal. Une autre option de sortie consiste à sauter le filtre inverse dans le programme de décompression. Un tel produit obtenu peut être traité de la manière ordinaire sans décon-
15 volution avant sommation. Pour la première option, la trace $[-2, -1, +1, +2]$ à quatre états est égalisée en gain échantillon par échantillon. Toutes les différences d'amplitude apparaissant à la sommation sont dues à la corrélation entre traces.

On se reporte maintenant aux figures 17 à 20 qui
20 donnent des descriptions détaillées des processus de création d'opérateur et de codage effectués avant la transmission et, ou bien, le stockage d'un signal comprimé, et des programmes de décondensation et de décodage effectués après la fin du stockage ou à la réception d'un signal comprimé. La figure 17 montre l'organigramme de la
25 création de l'opérateur. Une fois les échantillons entrant $S(t)$ reçus, une série d'échantillons d'une longueur prédéterminée représentative des données qui seront comprimées est choisie par calcul de l'autocorrélation. On retire la moyenne de la série afin d'éliminer tout niveau continu. On calcule ensuite l'autocovariance sur
30 les échantillons entre zéro et "Nmax" décalage temporels afin de déterminer la relation linéaire existant entre l'échantillon actuel et l'unité pour chaque décalage temporel. On localise ensuite la valeur absolue maximale entre "Nmin" et "Nmax", celle-ci étant la position de la longueur N_a de l'opérateur. On utilise ensuite
35 le décalage temporel zéro pour normaliser l'autocorrélation et fixer le côté droit des équations de Toeplitz citées ci-dessus

qui sont utilisées pour l'obtention de l'opérateur $A(n)$. Un certain pourcentage "Pcent" de bruit blanc est ajouté, et l'ordonnée centrale de l'autocovariance et l'opérateur $A(n)$ sont ensuite calculés par résolution des équations différentielles simultanées mentionnées

5 ci-dessus. S'il est trouvé au moyen de ce programme une solution stable aux équations, l'étiquette d'erreur est négative si bien que les valeurs comprises entre $N_a + 1$ et N_{max} peuvent être mises à zéro et que $A(n)$ peut être utilisé dans la compression. La longueur de l'opérateur $A(n)$ dépend de la longueur d'opérateur stable maximale,

10 c'est-à-dire de l'erreur. Si l'étiquette d'erreur est positive, ce qui indique que l'opérateur n'est pas stable, on ajoute 0,5% de bruit blanc à "Pcent" et on recommence le calcul. Si "Pcent" est supérieur à 10%, ce qui indique que le processus a été effectué 18 fois, aucun opérateur n'est calculé sur la base de cette série d'échantillons.

15 La figure 18 montre l'opération de codage effectuée après que le calculateur $A(n)$ a été calculé de la façon indiquée ci-dessus. On initialise d'abord le programme en calculant la valeur "Sig", qui est la moyenne des valeurs absolues des 25 premiers échantillons non nuls. On fixe ensuite les facteurs de dilatation

20 et de contraction d'après les articles de Jayant en fonction du nombre de bits de données codées voulus, et on choisit la taille depuis le quantum Q proportionnelle à "Sig", qui a été calculé de la manière indiquée ci-dessus. On met à zéro la valeur "Xpret", qui représente la valeur attendue suivante, et on met à un un registre de compteur

25 "Last". On calcule ensuite l'erreur E en utilisant l'opérateur tel que calculé ci-dessus, et celle-ci est égale à la valeur réelle $S(t)$ moins "Xpret". On quantifie la différence par comparaison de E avec les divers puits, ce qui fournit la valeur d'erreur quantifiée ET . Le numéro du puits à l'intérieur duquel l'erreur quantifiée s'ajuste,

30 soit $M(t)$, le numéro de quantum, est condensé et stocké pour servir de partie de l'enregistrement à transmettre. Le processus de quantification est détaillé ci-dessous en relation avec la figure 19. On vérifie ensuite la stabilité du processus en comparant la valeur absolue de l'échantillon avec l'intervalle de quantum, c'est-à-dire

35 avec la taille du puits. De façon claire, si la taille du puits est plus grande que la valeur absolue de l'échantillon, pour plusieurs

échantillons, une erreur s'est produite. Si la vérification de la stabilité révèle que l'opération se déroule de manière convenable, on peut alors calculer la valeur d'entrée approchée S à partir de l'équation $ET + X_{pret}$. Cette valeur peut alors subir la convolution avec $A(n)$ au cours de l'opération suivante de manière à produire la nouvelle valeur attendue " X_{pret} ". S'il s'agit du dernier échantillon, la trace suivante est lue, sinon, le processus recommence.

Comme cela a été mentionné ci-dessus, la figure 19 donne le détail du processus de quantification utilisé dans le processus de codage détaillé en relation avec la figure 18. L'erreur E doit être séparée en une taille de puits quantifiée et une erreur estimée, soit $M(t)$ et ET respectivement. Le processus commence par l'étalonnage de Q en fonction de la valeur "Last" disponible, c'est-à-dire que la taille de puits est fixée à une valeur de départ. La dimensions de puits ainsi déterminée est comparée avec " Q_{min} " et " Q_{max} ", c'est-à-dire avec les tailles maximale et minimale de puits précédemment déterminées, ceci constituant une vérification d'erreur. On utilise la valeur absolue de E dans les calculs qui suivent. On initialise les variables x et n , tandis que "Last" est fixé à une valeur égale à "ISTATE", qui vaut la moitié du nombre total de puits disponibles. Par exemple, dans le cas à 2 bits, quatre puits sont disponibles, car l'un des quatre puits peut être caractérisé par un code à 2 bits, à savoir : les petits échantillons positifs tombent dans le puits zéro; les grands échantillons positifs tombent dans le puits 1; le puits 2 est destiné aux petits échantillons négatifs; et le puits 3 aux grands échantillons négatifs. Dans ce cas, "ISTATE" vaut 2. La variable x est alors fixée égale à $x + Q$, et la valeur absolue de E est comparée avec x , ce qui revient à déterminer si E s'ajuste à l'intérieur de ce puits; si tel est le cas, on détermine alors $M(t)$. Enfin, le dernier puits choisi est choisi égal à n ; l'erreur quantifiée est choisie égale à la taille de puits $Q(\text{Last}-1/2)$. Si E est inférieur à zéro, on inverse ET , tandis que l'on choisit M égal à $\text{Last} + \text{ISTATE}$, c'est-à-dire au puits à l'intérieur duquel E échoit. Lorsque E est supérieur ou égal à x , on incrémente x de Q , c'est-à-dire la taille de puits, et le processus se répète. On note que ces dernières étapes sont rendues nécessaires par le fait que

l'on souhaite transmettre 2 bits sans signe. Ainsi, les numéros des puits correspondant aux valeurs centrées sur 0, c'est-à-dire correspondant à $0 < E \leq Q$, $E > Q$, $0 > E \geq -Q$, et $E < -Q$, sont choisis égaux respectivement à 0, 1, 2 et 3. Le transfert de "Last" à l'opération

5 "échantillonner Q", par laquelle le programme de quantification commence, permet de faire varier la taille du puits lorsque le processus se poursuit, ceci rendant la quantification souple. Par exemple, si le puits 1, c'est-à-dire le plus grand des puits, est celui à l'intérieur duquel tombe la valeur absolue de E, il faut

10 alors incrémenter la dimension de puits générale à l'aide de l'opération d'étalonnage. Si, d'autre part, la valeur absolue de E tombe dans le puits 0, il faut réduire la dimension de puits, car il est clair que E possède une petite valeur et, par conséquent, pour augmenter la précision, il faut rendre la taille de puits relative-

15 ment plus petite.

Enfin, les figures 20a et 20b montrent le processus de décondensation et de décodage effectué à la réception de données comprimées selon le procédé de l'invention en un lieu de réception ou à la lecture de ce qui a été stocké. Ainsi que cela a été noté

20 ci-dessus, on peut décoder les données pour obtenir une trace reconstruite ayant simplement subi la déconvolution ou ayant subi une égalisation en gain et une déconvolution, selon ce qui est choisi. Les variables entrantes sont des enregistrements de $M(t)$, les puits dans lesquels chaque échantillon d'erreur quantifiée

25 s'ajuste, de $A(n)$, l'opérateur, et de "Sig", la valeur initiale assignée à la dimension de puits. On met à zéro les échantillons initiaux et on fixe les facteurs de dilatation et de contraction de la manière indiquée ci-dessus. On initialise de même Q de la manière indiquée ci-dessus. On compare Q avec "Qmax" et "Qmin"

30 afin de déterminer la validité de cette valeur. La valeur prédite quantifiée est l'amplitude d'abord donnée par le numéro de puits et le niveau de quantification Q, puis le signe est appliqué.

Si "Tout", qui est une variable indiquant le mode voulu de traitement des données, est égal à 2, ce qui indique que

35 les données comprimées stockées doivent subir une égalisation en gain et une déconvolution, $S(t)$ est simplement choisi égal à

"Last" si M vaut de moins que "ISTATE", et à "-Last" dans le cas contraire, c'est-à-dire que la fonction de sortie est simplement égale à l'amplitude relative et au signe de l'échantillon d'erreur. Il se révèle que, en général, même ceci produit un résultat significatif. Si "Iout" n'est pas égal à 2, ceci signifiant que la trace doit subir une déconvolution ou bien être reconstruite, un traitement supplémentaire est effectué. Si M est supérieur à "ISTATE", le signe de "ET" est négatif. Si "Iout" est égal à 1, la trace subit une déconvolution seulement, ce qui implique qu'on fixe simplement $S(t)$ égal à "ET", sinon, on choisit alors $S(t)$ égal au résultat de l'équation " $X_{pret} + E(t)$ ", ce qui produit un point de trace complètement reconstruit; " X_{pret} " est égal à la somme de $A(n) \cdot S(t-n)$ qui produit la valeur attendue suivante " X_{pret} ", et le processus se répète.

15 Sur la base des organigrammes des figures 17 à 20, l'homme de l'art se trouvera facilement en mesure de mettre complètement en oeuvre le procédé de l'invention.

L'homme de l'art notera donc qu'il a été décrit un procédé de compression de données s'utilisant spécialement dans les applications d'exploration sismique qui autorisent la représentation d'échantillons de 32 bits envirgule flottante par un entier de 2 bits sans perte de fidélité de reproduction. L'utilisation du concept de quantification souple et de modulation par impulsions codées différentielles est seule insuffisante pour donner le résultat, car le bruit résultant d'un tel schéma lorsque ce dernier est appliqué à des données sismiques conduit à des données instables sans signification. Par conséquent, le signal est transformé en une trace d'erreur de prédiction, au lieu d'une trace de données, avant la quantification; selon une caractéristique particulièrement préférée, on obtient ce résultat en utilisant la méthode de modulation par impulsions codées différentielles à prédiction linéaire, modifiée de la manière indiquée ci-dessus.

Bien entendu, l'homme de l'art sera en mesure d'imaginer, à partir du procédé dont la description vient d'être donnée à titre simplement illustratif et nullement limitatif, diverses variantes et modifications ne sortant pas du cadre de l'invention.

REVENDICATIONS

1. Procédé d'exploration sismique dans lequel de l'énergie acoustique (12) est transmise au travers d'une structure souterraine (16), l'énergie réfléchie par une interface souterraine (18) et renvoyée vers le haut jusqu'à la surface est mesurée (20), des
5 signaux électriques de données de sortie (22) sont produits en réponse à l'énergie ainsi mesurée, et les signaux de données sont comprimés (24, 28, 30) en vue d'un stockage et, ou bien, d'une transmission (32, 34, 36, 38), l'opération de compression des données consistant à :
- échantillonner les signaux à un taux fixe;
10 analyser plusieurs séries d'échantillons de sortie successivement en fonction d'un opérateur de prédiction de manière à produire une valeur prédite pour l'échantillon d'entrée suivant;
comparer l'échantillon d'entrée suivant avec la valeur prédite correspondante afin de produire une valeur d'erreur
15 de prédiction;
quantifier la valeur d'erreur de prédiction de manière à produire un nombre de quantum; et
stocker et, ou bien, transmettre le nombre de quantum;
- 20 l'opération de compression de données étant telle que, au moment de l'inversion de l'opérateur destinée à effectuer la décompression des nombres de quantum stockés, une réaction peut être appliquée à l'opérateur inversé afin de faire que le bruit se trouvant dans les échantillons quantifiés passe sans filtrage.
- 25 2. Procédé selon la revendication 1, caractérisé en ce que l'opérateur de prédiction comprend une modulation par impulsions codées différentielles à prédiction linéaire dans laquelle on utilise un nombre prédéterminé d'échantillons successifs pour prédire l'échantillon immédiatement suivant, et l'échantillon immédiatement suivant
30 devient ensuite le dernier des échantillons du nombre prédéterminé d'échantillons, le plus ancien des échantillons du nombre prédéterminé d'échantillons étant rejeté du nombre prédéterminé d'échantillons.
3. Procédé selon la revendication 1 ou 2, caractérisé en ce que la quantification est souple de sorte qu'on peut obtenir
35 une précision accrue dans la quantification des échantillons.

4. Procédé selon la revendication 1, 2 ou 3, caractérisé en ce que l'opération de quantification consiste à catégoriser chaque échantillon comme appartenant à l'un d'un nombre fixe d'états distincts, les états formant un intervalle continu dans lequel ceux des 5 états qui sont disposés au centre de l'intervalle ont des limites fixes et ceux des états se trouvant aux extrémités de l'intervalle ont une limite fixe pour concorder avec ceux du centre de l'intervalle et une limite ouverte, si bien que toutes les valeurs d'échantillons peuvent être catégorisées comme appartenant à l'un seulement des 10 états.

5. Procédé selon la revendication 4, caractérisé en ce que les limites des états appartenant au centre de l'intervalle varient en fonction de la catégorisation d'un ou de plusieurs échantillons antérieurs si bien que l'on peut encore améliorer la précision de la quantification. 15

6. Procédé selon l'une quelconque des revendications 1 à 5, caractérisé en ce qu'une série d'échantillons quantifiés correspondant à une période donnée d'exploration sismique est précédée d'un entête donnant l'information concernant les conditions initiales 20 du procédé de compression de données, de sorte que la décompression des données puisse être effectuée.

7. Procédé selon l'une quelconque des revendications 1 à 6, caractérisé en ce que l'opération d'analyse consiste à réaliser l'autoconvariance sur chaque série d'échantillons, de sorte qu'un 25 glissement de basse fréquence de la compression est évité.

8. Procédé de compression d'échantillons de signaux de nature convolvable, caractérisé en ce qu'il comprend les opérations suivantes :

définir un opérateur de prédiction par une analyse 30 effectuée relativement à un nombre prédéterminé des échantillons visant à agir sur une série des échantillons de longueur prédéterminée pour produire une valeur prédite pour un échantillon suivant;

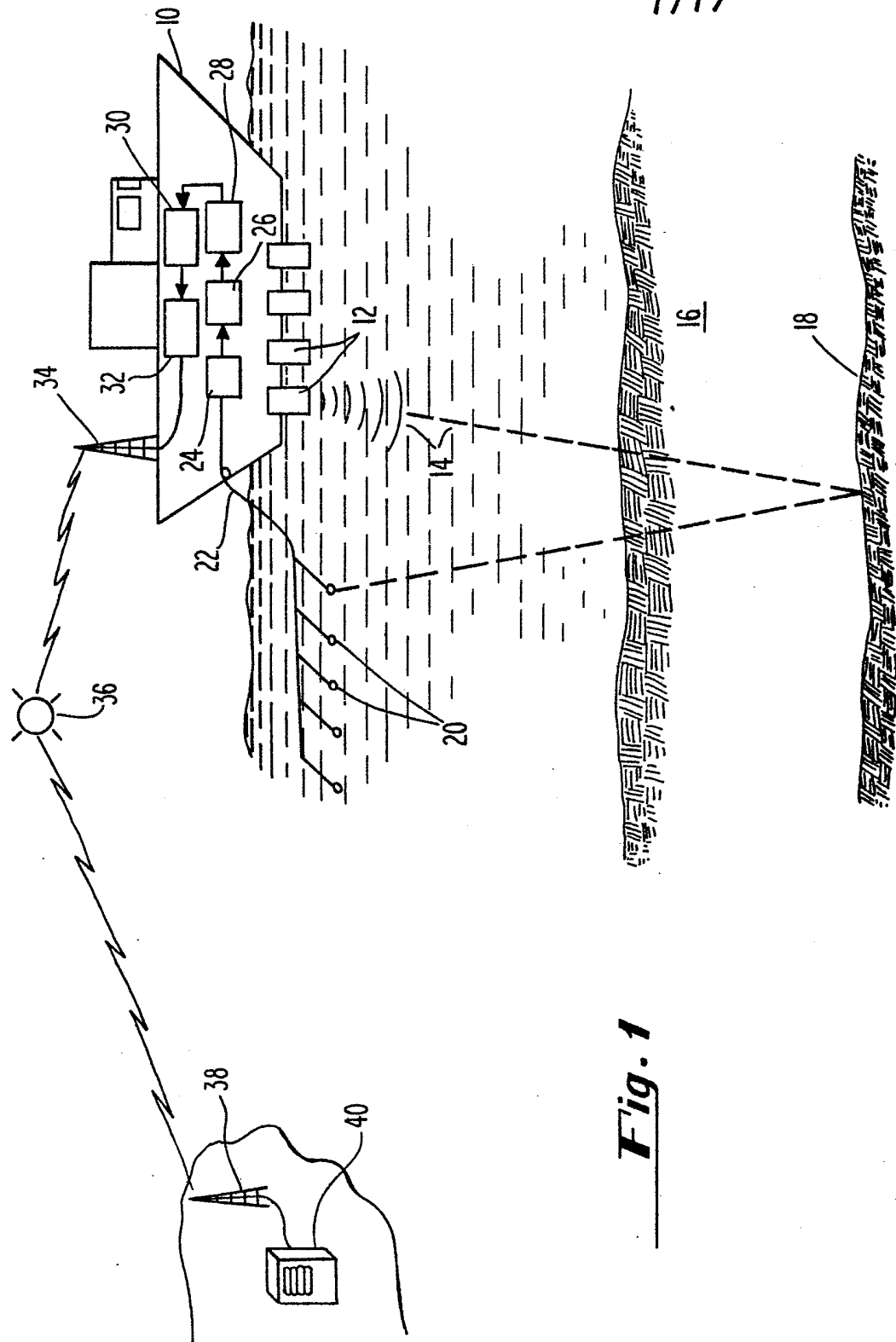
mesurer la différence existant entre la valeur prédite et l'échantillon suivant;

35 comparer la différence à une série de puits numérotés de dimension variable;

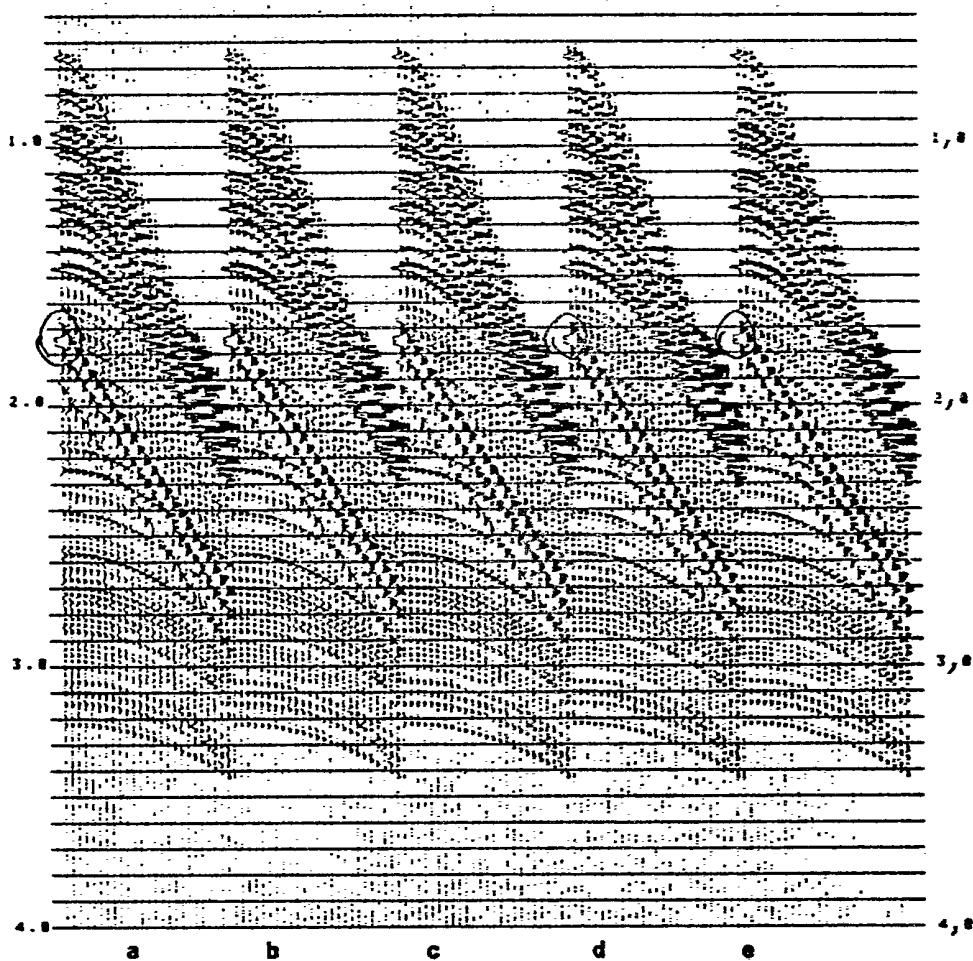
enregistrer le numéro du puits dans lequel la différence s'ajuste; et

faire varier la dimension des puits relativement au numéro.

1/17

**Fig. 1**

2/17

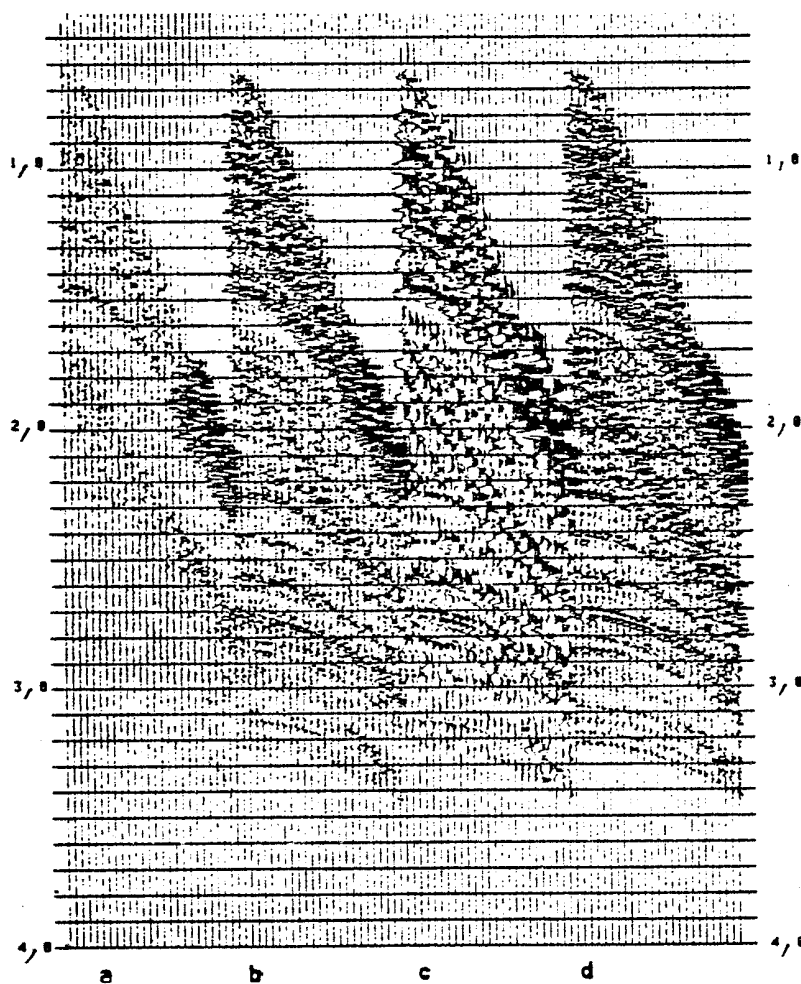


Cinq enregistrements d'essai en entrée

- a) codage d'un mot complet (FWC)
- b) DPCM à prédiction linéaire (LPDPCM)
- c) DPCM à somme courante (RSDPCM)
- d) DPCM simple (SDPCM)
- e) Codage du seul bit de gain (GBOC)

FIG. 2

3/17

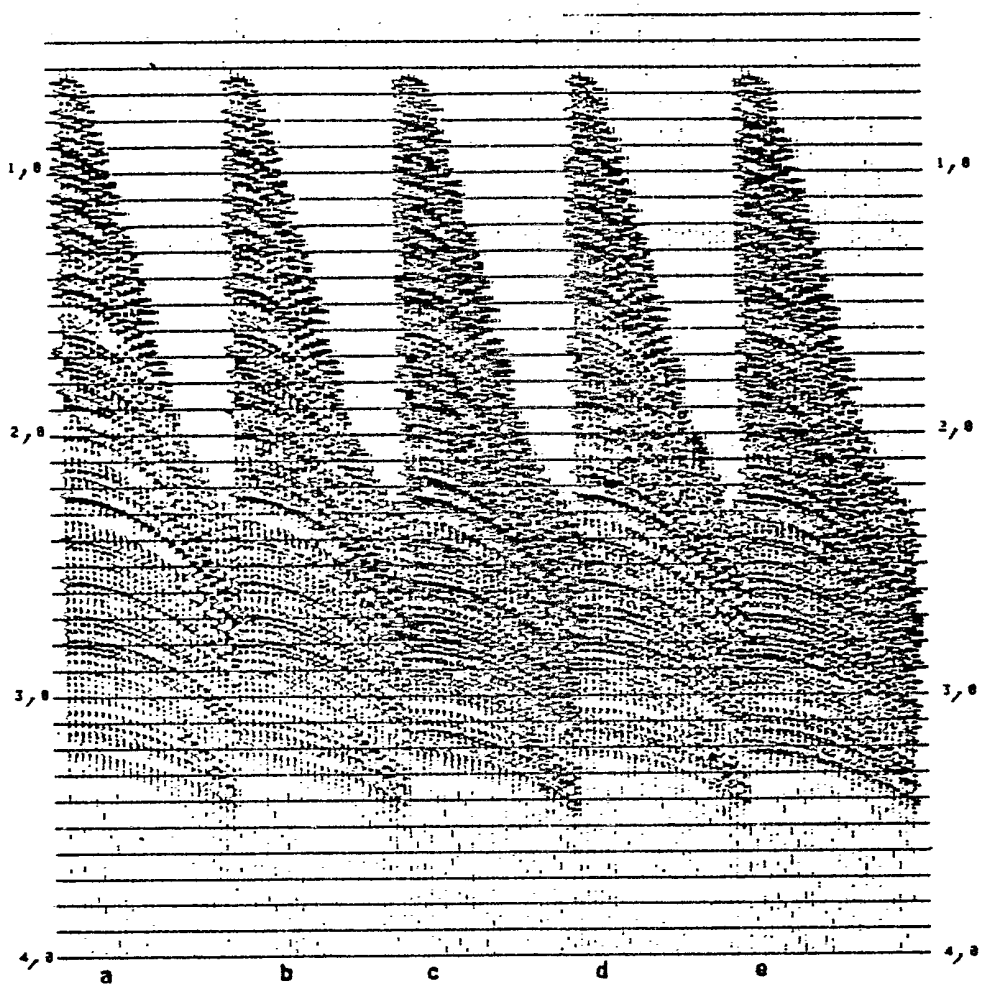


Enregistrements de différence en entrée

- a) FWC — LPDPCM
- b) FWC — RSDPCM
- c) FWC — SDPCM
- d) FWC — GBOC

FIG. 3

4/17

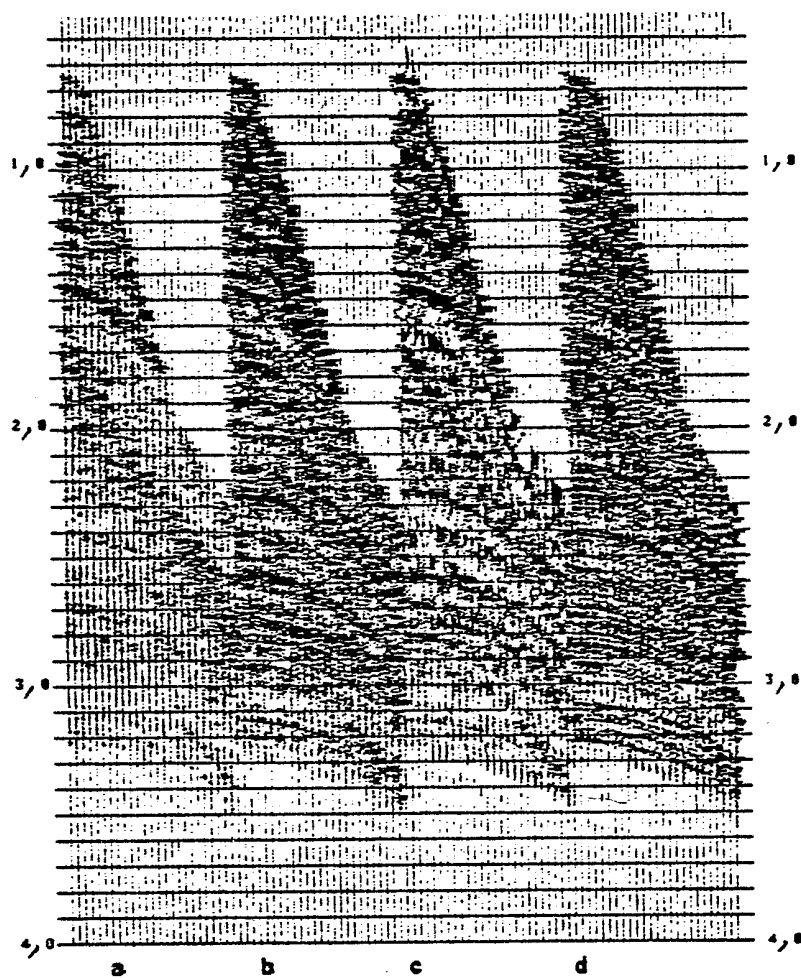


cinq enregistrements d'essai après décodage $f-k$ et
déconvolution de pointe

- a) FWC
- b) LPDPCM
- c) RSDPCM
- d) SDPCM
- e) GBOD

FIG. 4

5/17

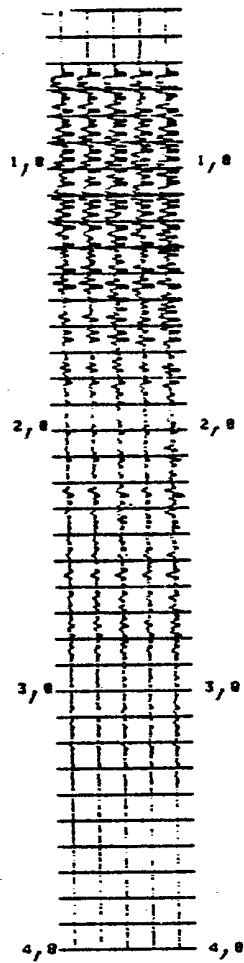


*Enregistrements de différence après filtrage f-R et
déconvolution de pointe*

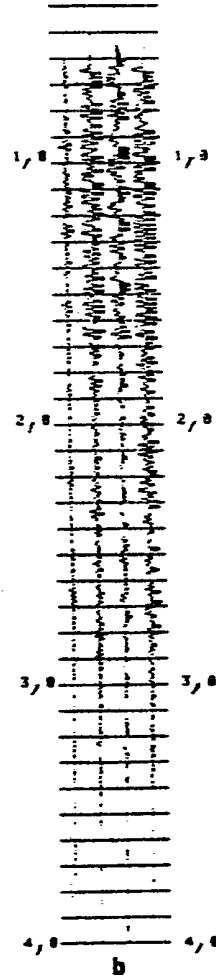
- a) FWC - LPDPCM
- b) FWC - RSDPCM
- c) FWC - SDPCM
- d) FWC - GBOC

FIG. 5

6/17

**FIG. 6**

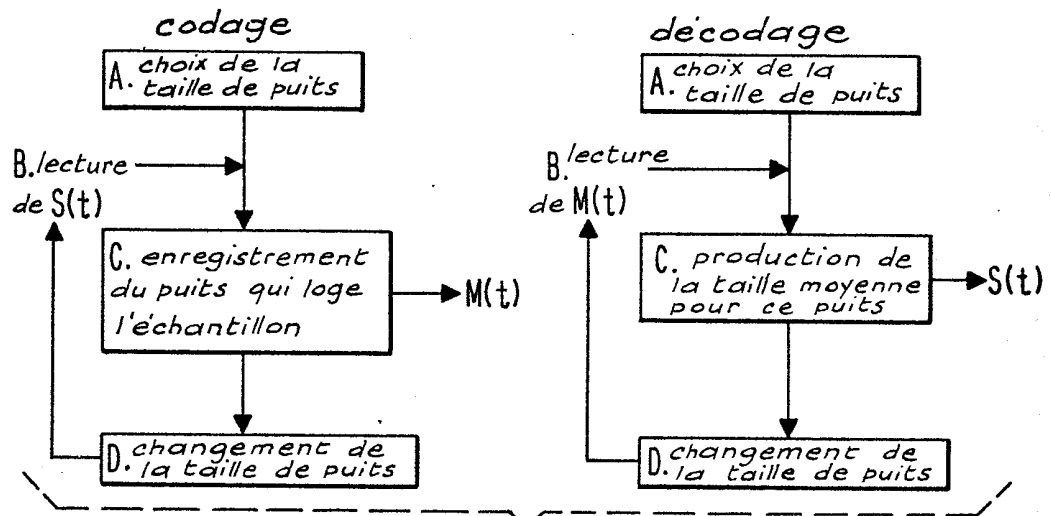
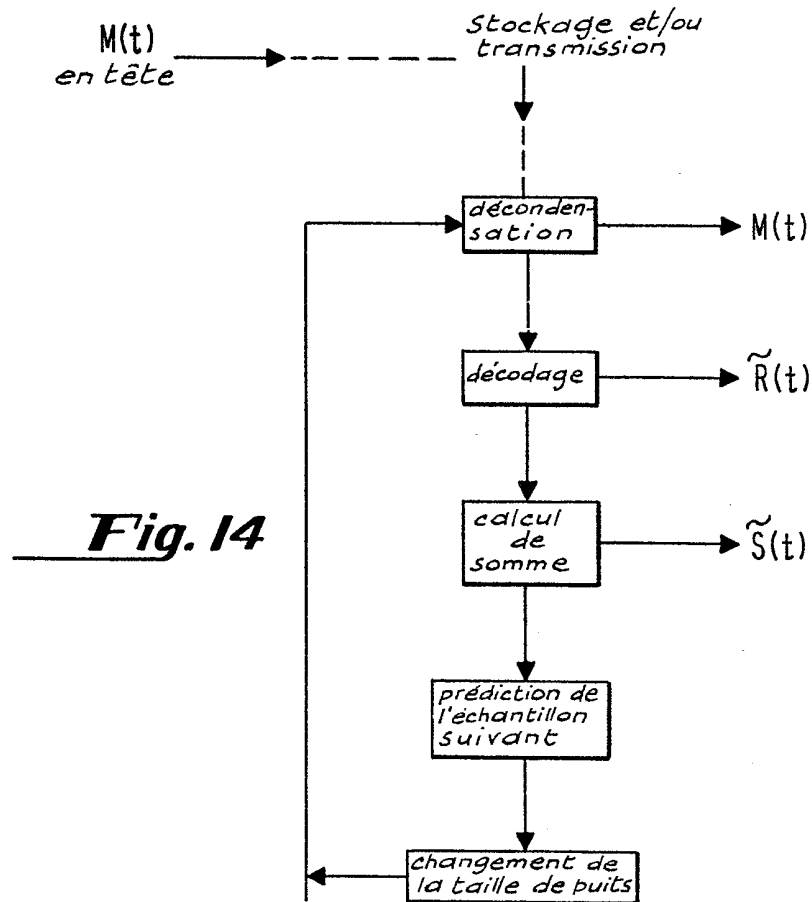
*cinq enregistrements
d'essai après corrections
dynamiques et sommation*

**FIG. 7**

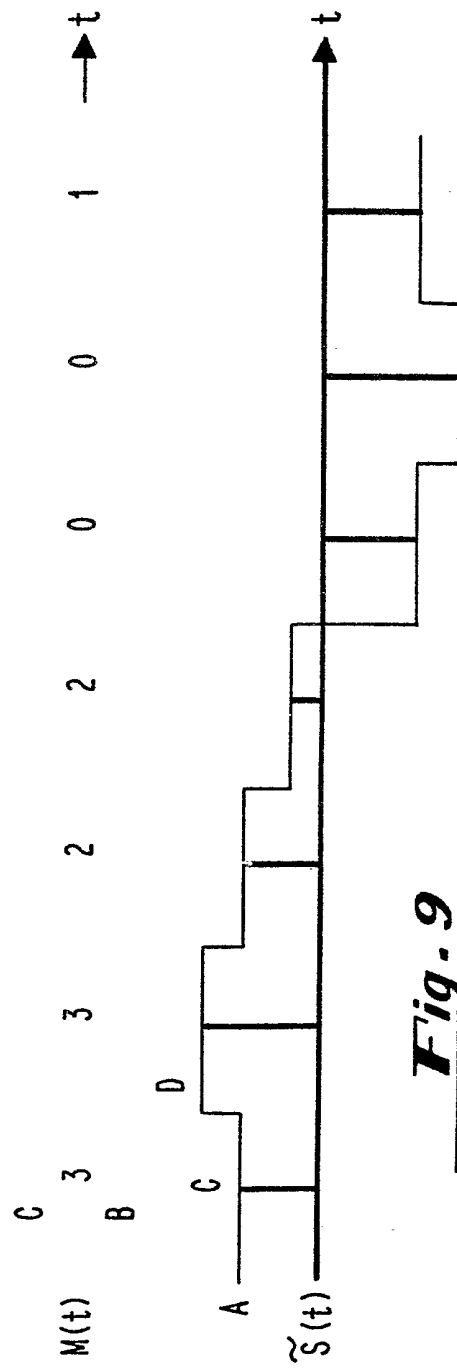
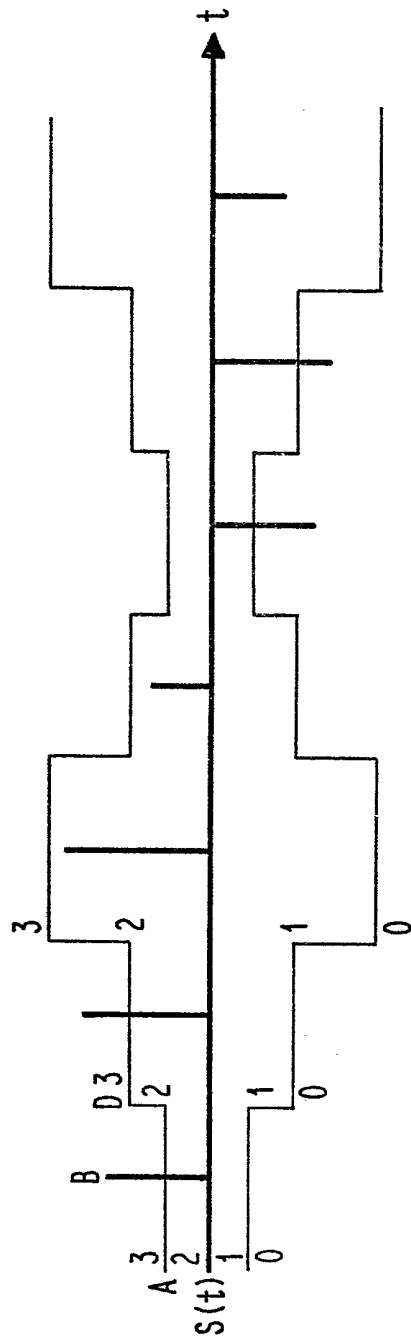
*Traces de
différences*

- 1) FWC - LPDPCM
- 2) FWC - RSDPCM
- 3) FWC - SDPCM
- 4) FWC - GBOC

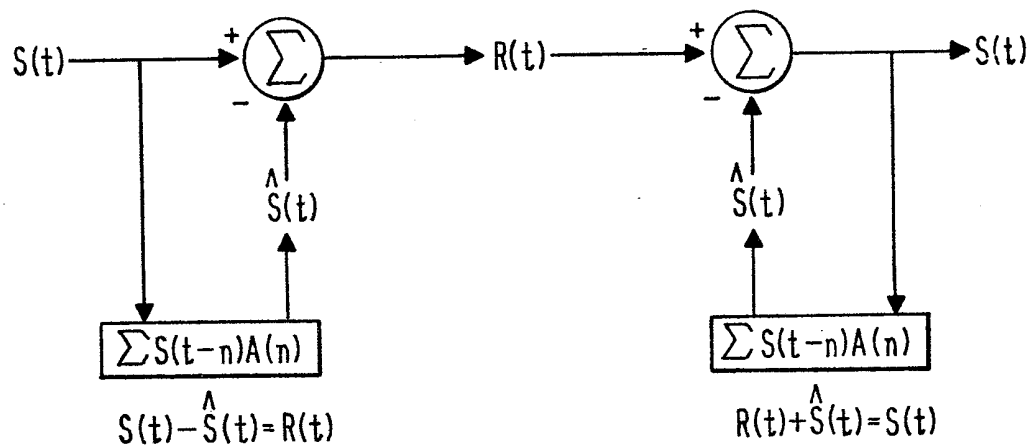
7/17

**Fig. 8****Fig. 14**

8/17

***Fig. 9***

9/17

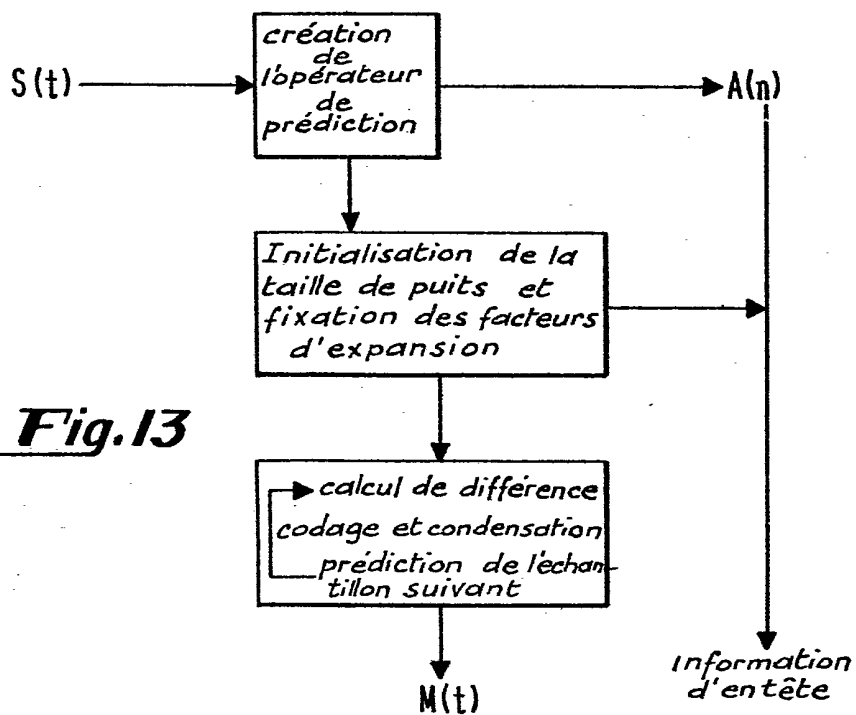
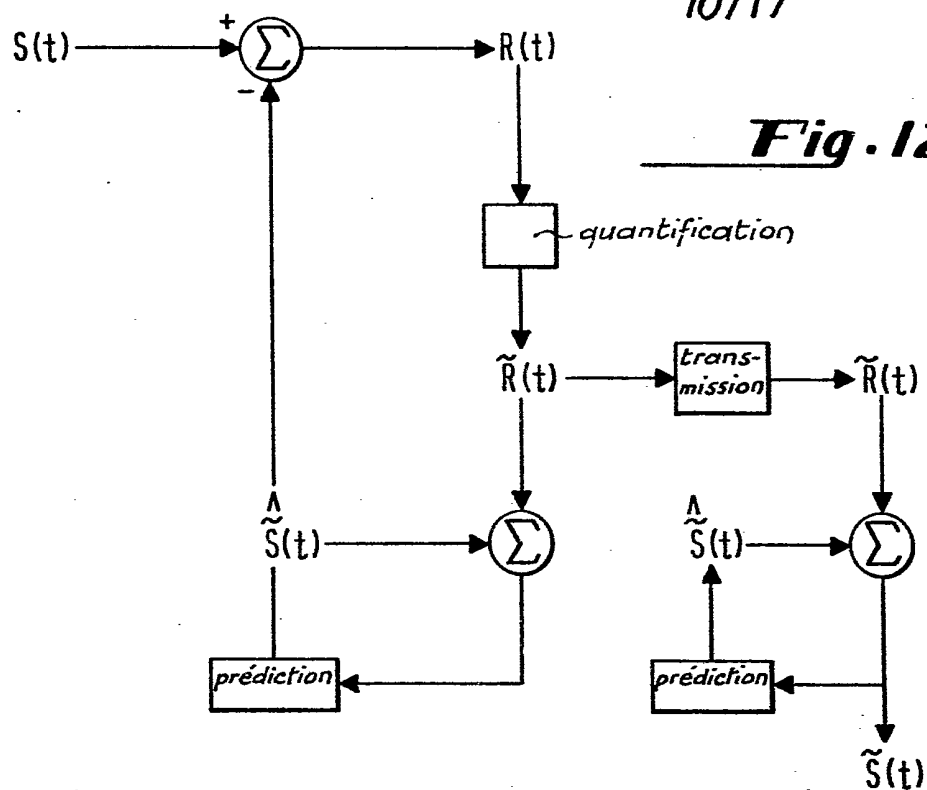
**Fig. 10**

$$\begin{aligned}
 I &= \sum R^2(t) \\
 &\downarrow \\
 R(t) &= S(t) - \hat{S}(t) \\
 &\downarrow \\
 \hat{S}(t) &= \sum S(t-n) A(n)
 \end{aligned}$$

Résoudre n équations pour l'opérateur $A(n)$
telles que $\frac{\partial I}{\partial A_n} = 0$ minimise I

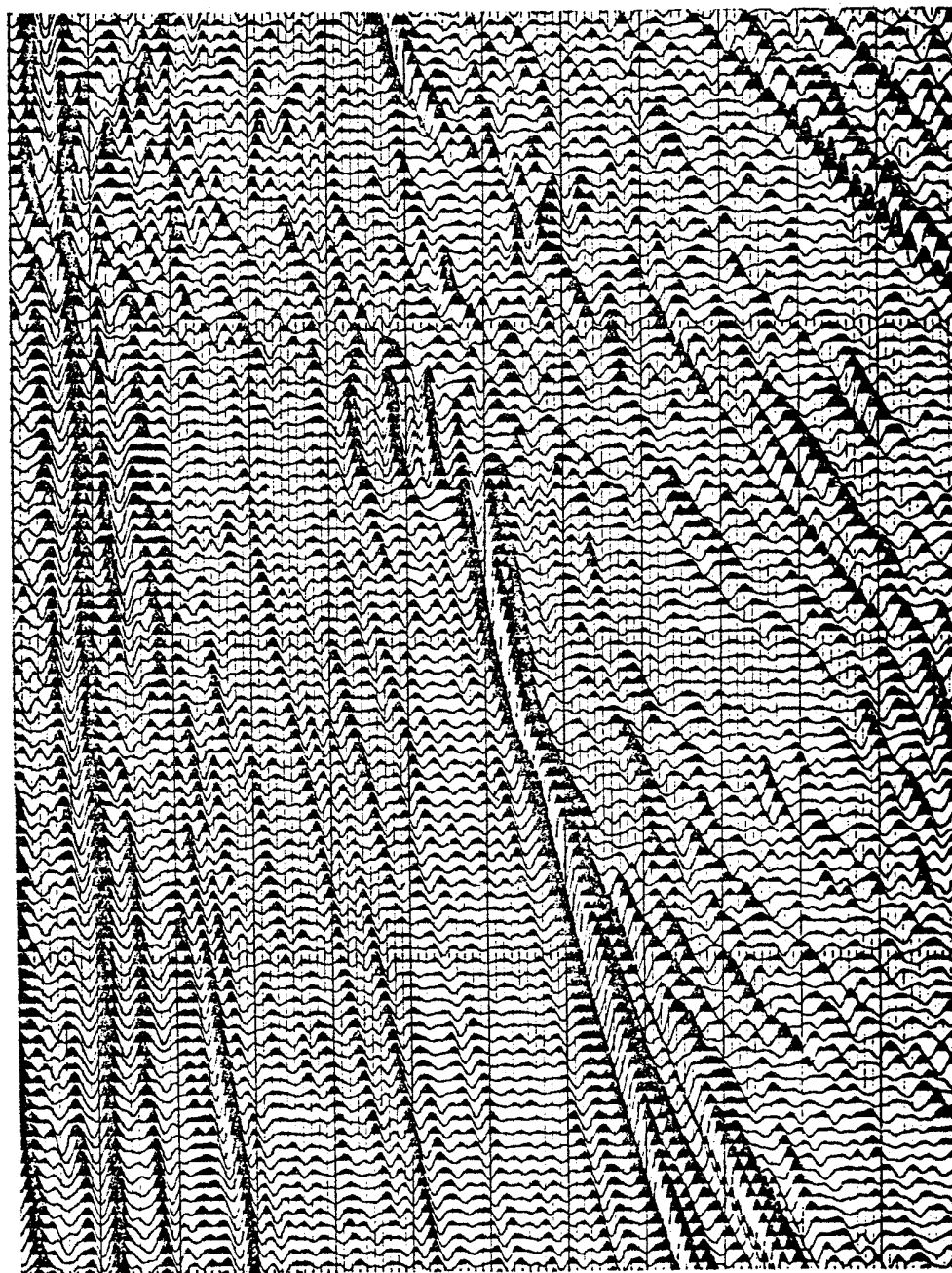
Fig. 11

10/17

Fig. 12

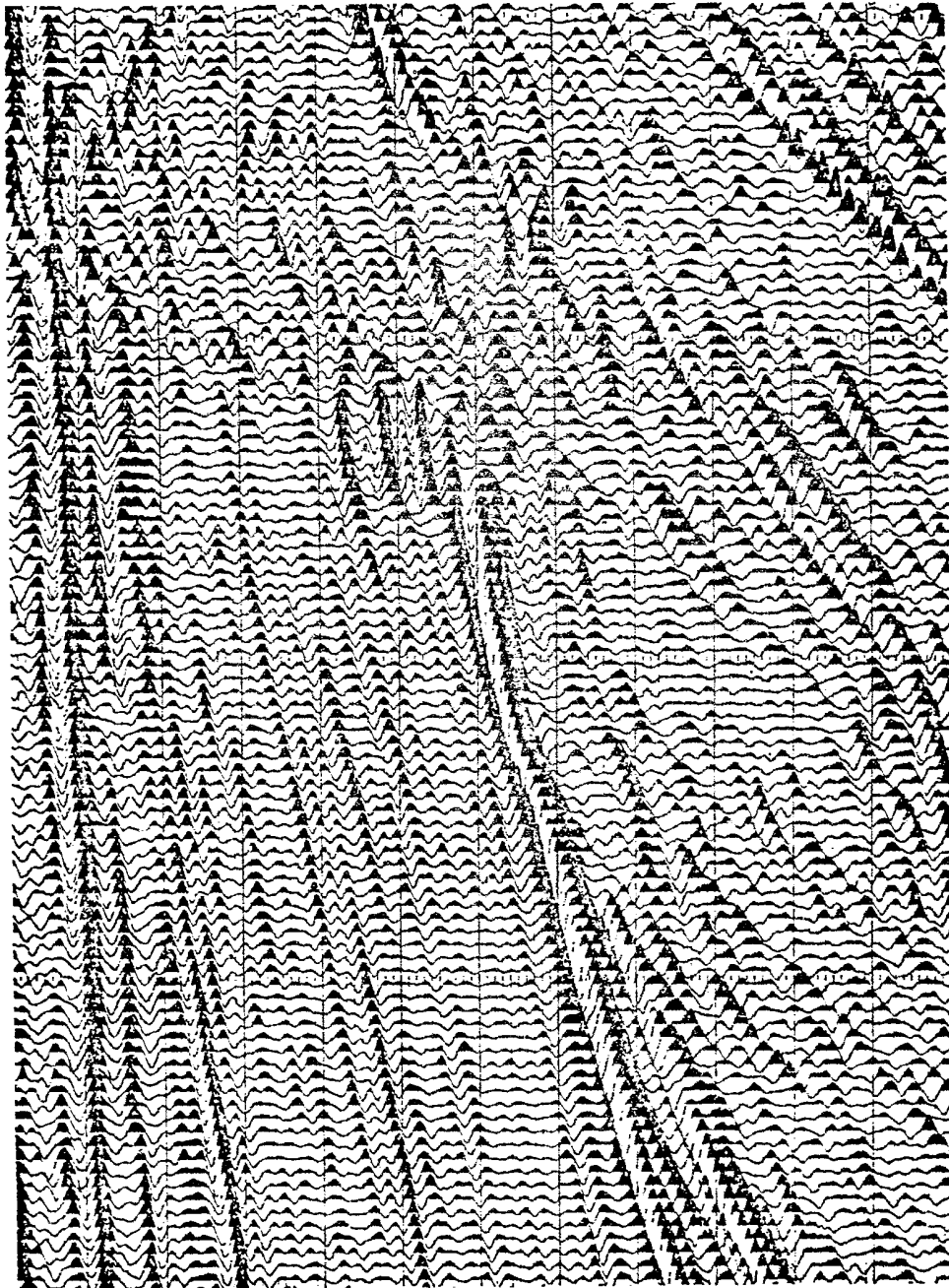
11/17

FIG.15 *Codage d'un mot complet*

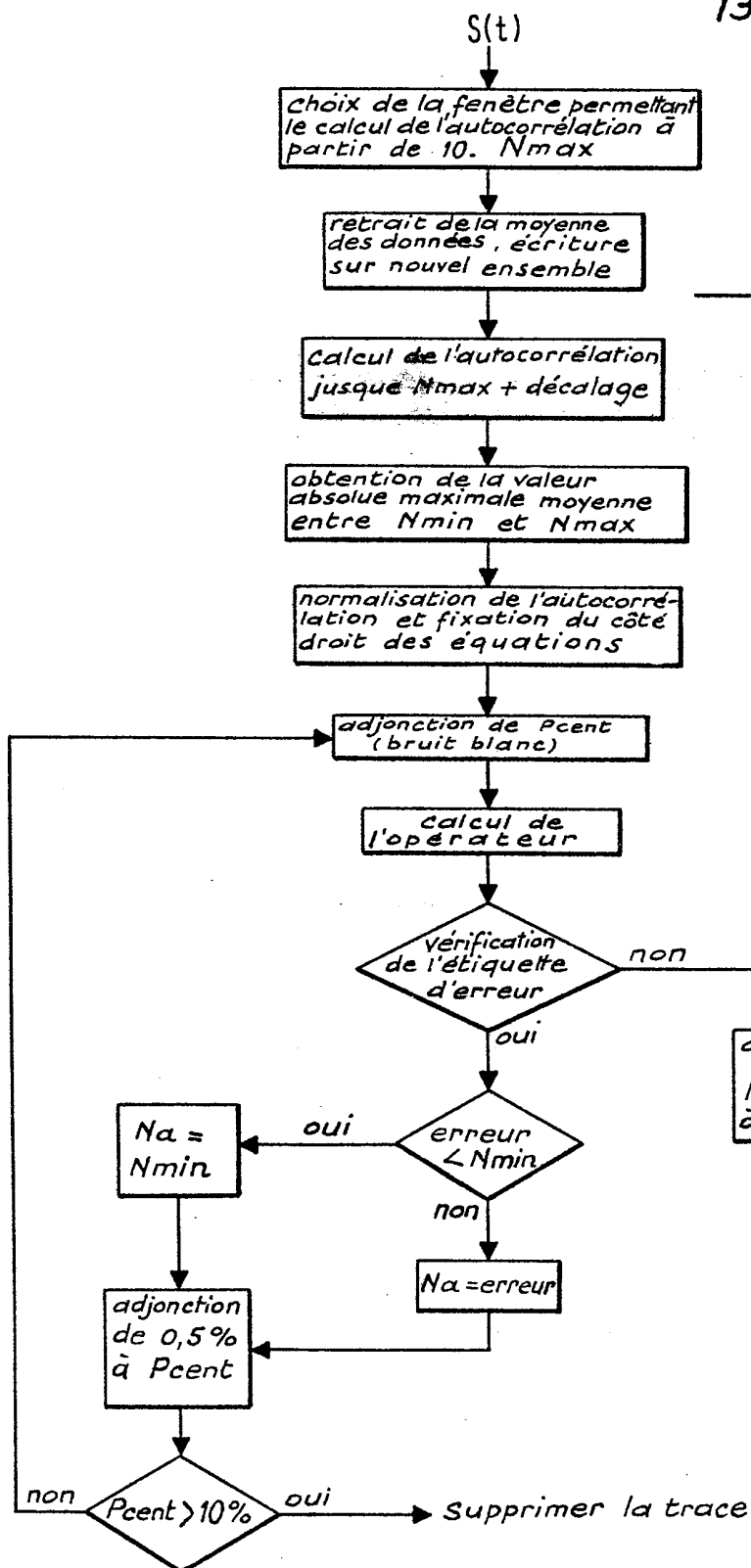


12/17

FIG.16 *Codage à deux bits par mot de décompression*

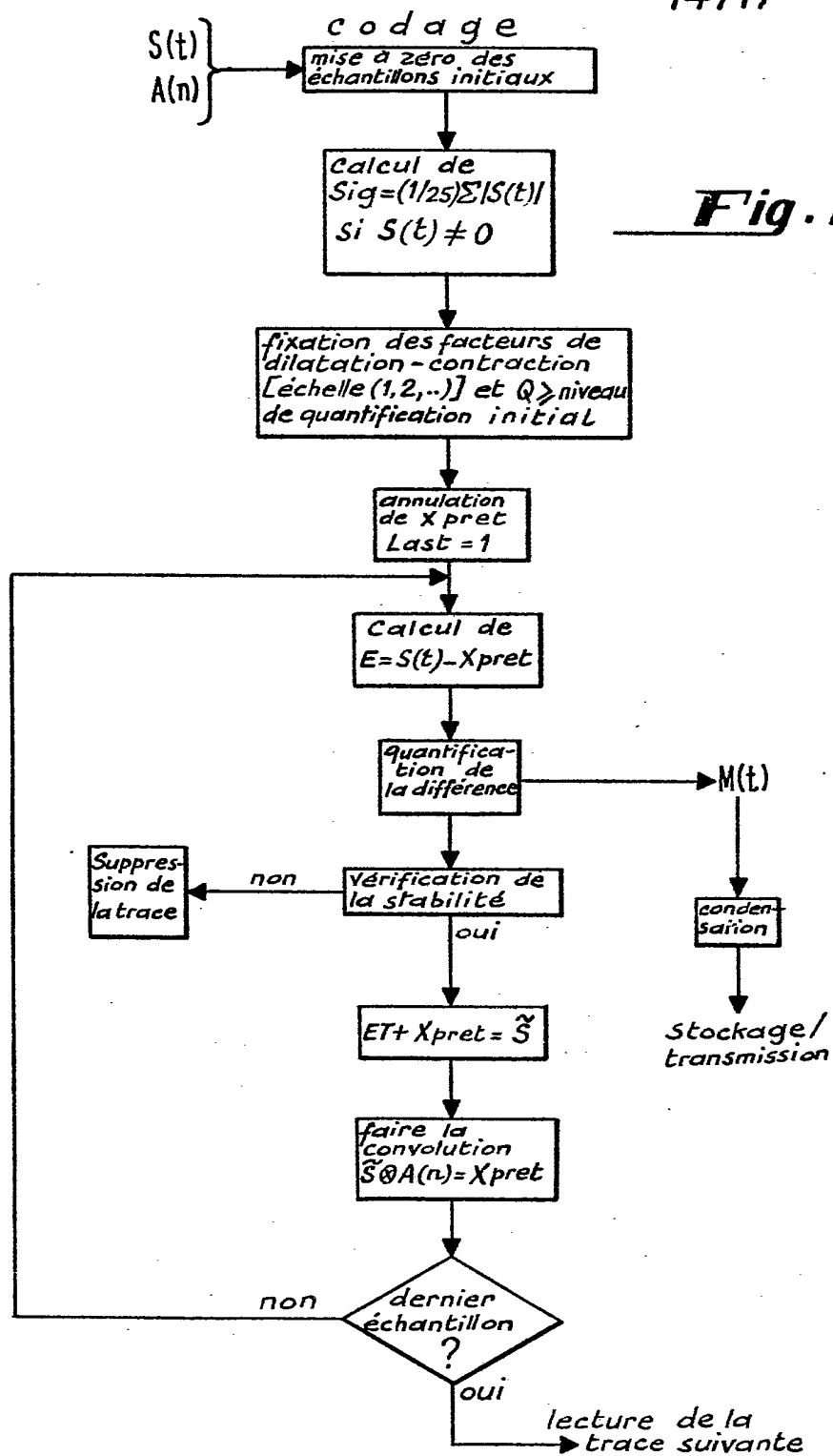


13/17

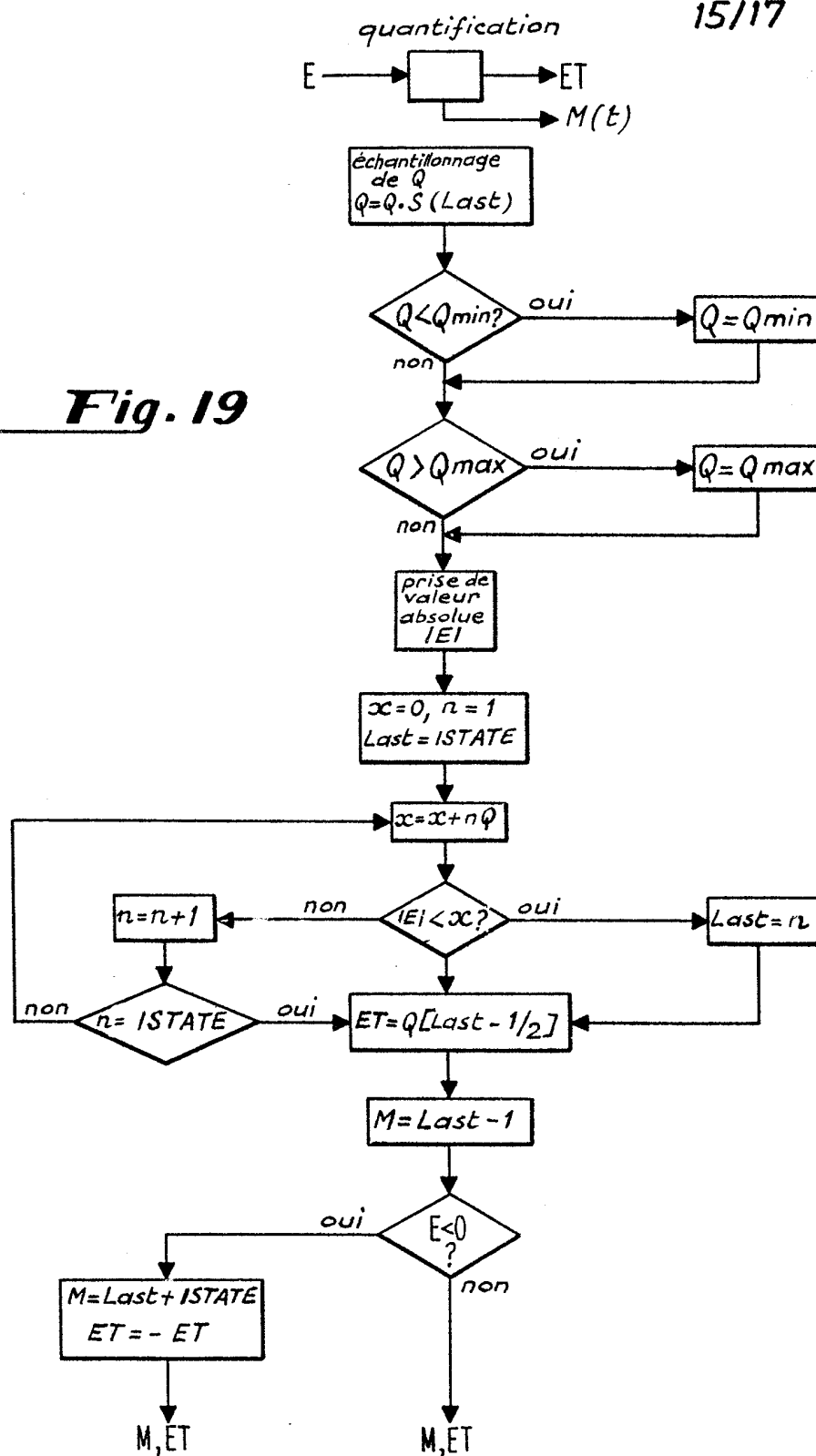
Fig. 17

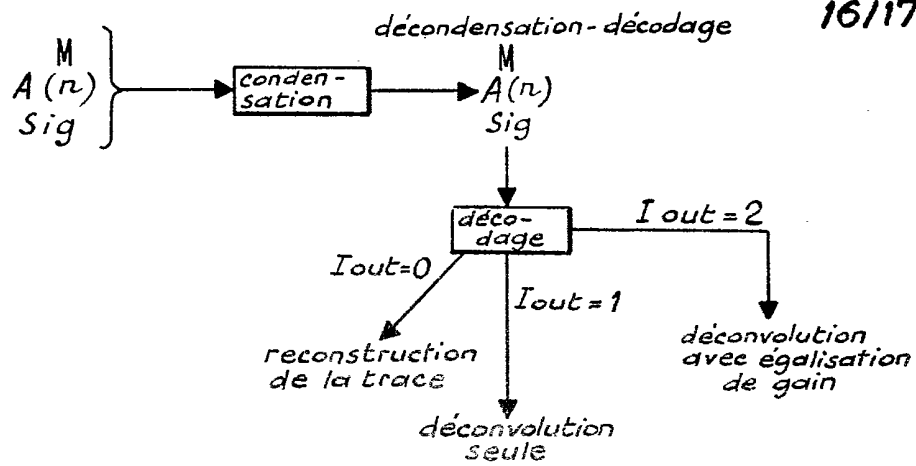
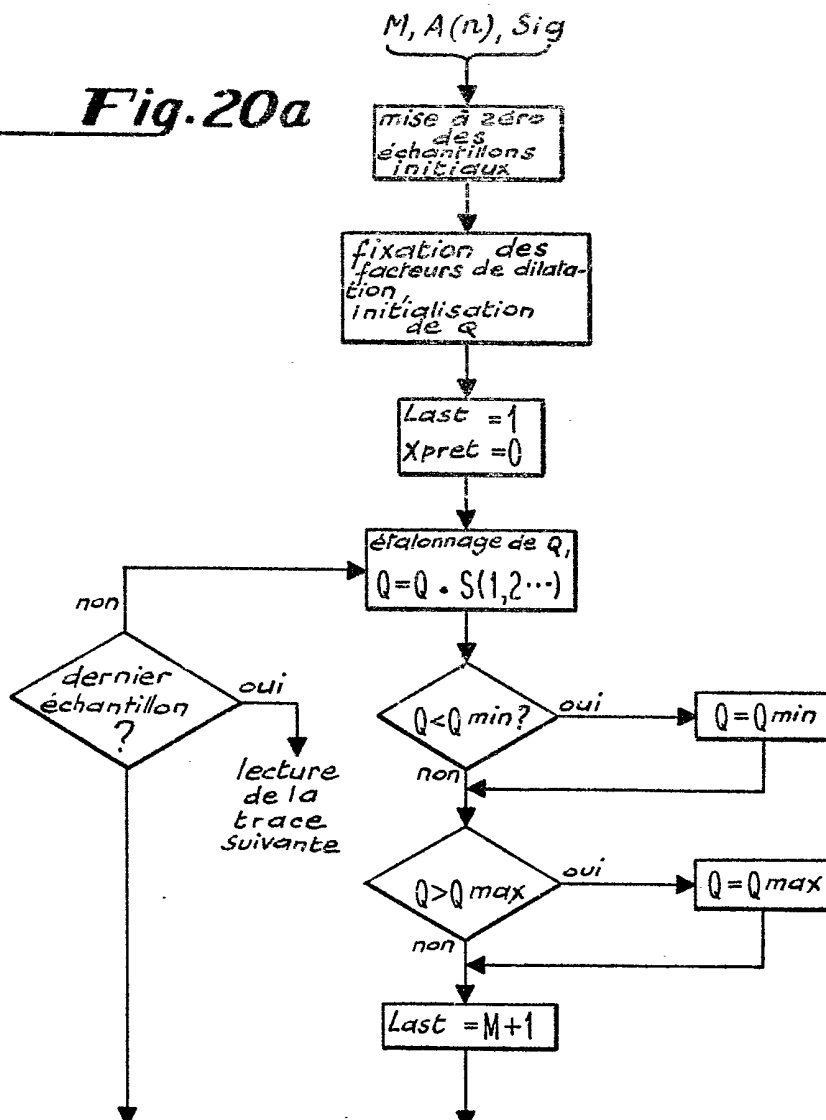
14/17

Fig. 18



15/17

Fig. 19

**Fig. 20a**

17/17

