

オーディオにおける デジタル技術

山崎 芳 男

ま え が き

放送や録音の分野においてデジタル技術は従来から制御系には使われていたが、最近では信号系への導入が盛んに検討され、一部では実用に供されている。BBCでは1972年以来音声にPCM中継を導入⁽¹⁾、すでにほぼ全国を網羅している。また、日本コロムビアではPCMレコーダをマスターレコーダに使ったレコードが1971年以来製作されている^{(5),(6)}。

ところで、信号伝送系のデジタル化の特徴は信号のSN比がアナログ伝送では中継器間隔と中継回数、伝送路等によって決るのに対し、デジタル伝送では量子化ステップ数だけによって決るという点であろう。すなわちデジタル伝送においては0,1の判断さえ間違いなく出来る伝送系であれば、信号劣化がないということである。またデジタル伝送機器は設置してしまえば、保守の点でも有利なはずである。

以上のようにデジタル伝送は、数多くの利点を有するが、アナログ伝送とは比較にならない程の伝送帯域を必要とするので実用化にあたっては多くの問題がある。一方、アナログ伝送には使えないような低品質の伝送路の使用が可能なので、衛生通信などには早くから使われ

ていた。

近い将来放送、録音等のオーディオ系の全てがデジタル化されることはないであろうが明らかにデジタル化した方が有利な機能、デジタル化しなくては実現できない機能はデジタル化され、アナログ系と併存していくものと思われる。

放送におけるデジタル技術、PCMの歴史等については、すでに本誌1974年月号のデジタルテレビジョンに林宏之氏が書かれているので本稿ではPCMの基本原則、各種機能のデジタル化について紹介をする。

1. PCMの原理

デジタル化とはアナログ量を離散的な値で表現することである。

PCMはいわばこのデジタル化を、アナログ信号の時間と振幅の双方に適用した伝送方式である。すなわち、等時間間隔毎に信号の振幅を読み出し、標本化し、その値を離散的な値で表現量子化し、さらにこの値をパルス符号で表す符号化するわけである。

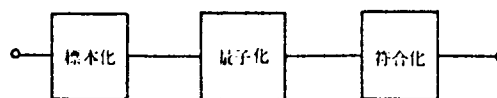


図1 PCMの基本原則

山崎芳男：早稲田大学・理工学部

1.1 標本化

標本化とは連続信号をある時点の値で代表させる操作である。標本化定理によれば関数の振幅を正確に定義するには、ある一定の離散的な時間間隔で標本化を行えば十分である。すなわち、

「振幅と時間の関数である信号が等間隔で少なくとも信号の最高周波数の2倍の速度で標本化されるなら、その標本列にもとの信号の全情報が含まれる。」からである。

$\frac{1}{2T}$ (Hz) に帯域制限された信号を $g(t)$ とすると、 $g(t)$ は標本化定理により、

$$g(t) = \sum_{n=-\infty}^{+\infty} g(nT) \cdot \frac{\sin \frac{\pi}{T}(t-nT)}{\pi(t-nT)} \quad \dots\dots\dots (1)$$

と書ける。右辺第2項は帯域幅 $\frac{1}{2T}$ の理想低域フィルタのインパルス応答に相当する。従って式は原信号 $g(t)$

が 秒間隔の標本値の全てを帯域 $\frac{1}{2T}$ の理想低域フィルタに通すことにより再現できることを示している。

実際には完全な帯域制限、幅のないパルスによる標本化、幅のないパルスと理想低域フィルタによる復調は実現不可能である。

帯域制限が不完全であると、標本化したスペクトラムのベースバンドと最初の側帯波に重りが生ずる。この不要スペクトラムは当然信号と相関があり折り返しひずみとよばれる。

一方、有限幅パルスで標本化、復調を行うと周波数ひずみが発生する。標本化周期を T (秒)、標本化パルス幅を τ (秒) とすると、標本値は、

$$h(t) = \frac{\tau}{T} \cdot \frac{\sin \frac{\tau}{2} \omega}{\frac{\omega}{2}} g\left(t - \frac{\tau}{2}\right) \quad \dots\dots\dots (2)$$

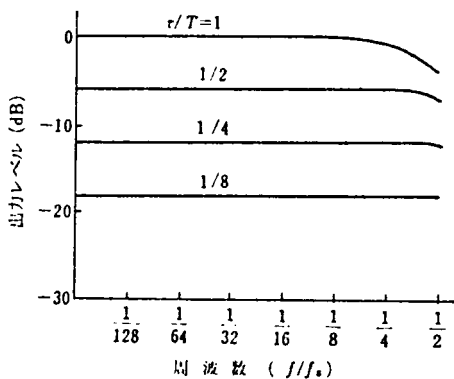


図2 保持効果

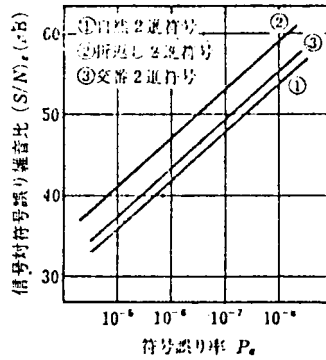


図3 符号誤りによる S/N と通話路当たり雑音量

となり、 $\tau/2$ の時間遅れと $\frac{\tau}{T} \cdot \frac{\sin \frac{\tau}{2} \omega}{\frac{\omega}{2}}$ の周波数ひ

ずみが生ずる。これを保持効果とよぶ。時間遅れはひずみを発生しないが、周波数特性は τ/T により図2のような変化を受ける。

広帯域音響信号の伝送にあたっては、帯域を 20kHz とすると標本化周波数は 40kHz 以上必要となる。一方、保持効果は標本化過程においてはサンプルホールド回路のパルス幅は一般に、十分細く問題ないが、復調時にはSN比の問題から τ/T をあまり小さくすることはできない。SN比と周波数特性の兼ね合いから $\tau/T = 1/4 \sim$

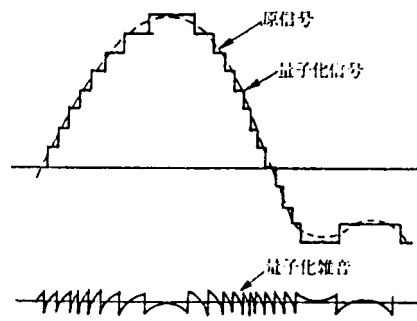


図4 量子化と量子化雑音

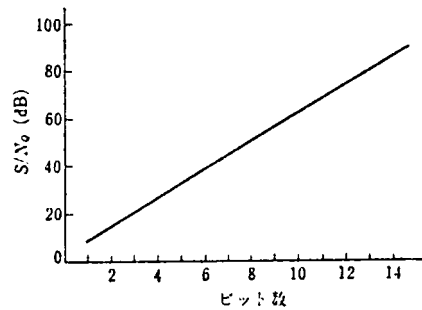


図5 量子化雑音

1/10 が適当といえよう。

1.2 量子化・符号化

標本値を離散的ないくつかの振幅で表現することを量子化という。そしてこの振幅値をパルス符号に変換するわけであるが、一般には2進符号が用いられ、その桁数がビット(Binary Digit)とよばれる。8ビットでは、 $2^8 = 256$ 、12ビットでは $2^{12} = 4096$ の振幅値が表現できる。2進符号には自然2進、折り返し2進、交番2進等がある。伝送路に誤りがあるとき各符号系では若干SN比が異なり、折り返し2進が有利である(図3)。

ところで量子化は標本値を離散的な値で表現する操作であるから、丸めによる誤差を生ずる。この誤差を量子化雑音とよび、PCMが原理的に有する雑音である(図4)。一様量子化による雑音電力 N_q は、量子化ステップを ΔV とすると、

$$N_q = \frac{(\Delta V)^2}{12} \dots\dots\dots (3)$$

として与えられる。従って最大振幅の正弦波と量子化雑音のSN比は、ビット数を n とすると、

$$S/N_q = \frac{3}{2} \cdot 2^n = 6n + 1.8(\text{dB}) \dots\dots\dots (4)$$

「広帯域音響信号の伝送においてダイナミックレンジは70~80dBは必要である。」(図5) 従って、一様量子化では12~13ビット必要である。

量子化雑音はトランジスタや抵抗の熱雑音、テープヒス等とは違い、信号が存在するときだけ発生する一種のひずみであるが、量子化ステップ数が多いとスペクトラムは白色に近づく。しかし、最下位ビット(LSB)にかかるかからない程度の小レベル信号は、大きなひずみ(granular distortion)を発生する。このひずみを減らすために“dither”とよばれる小振幅の白色雑音や正弦

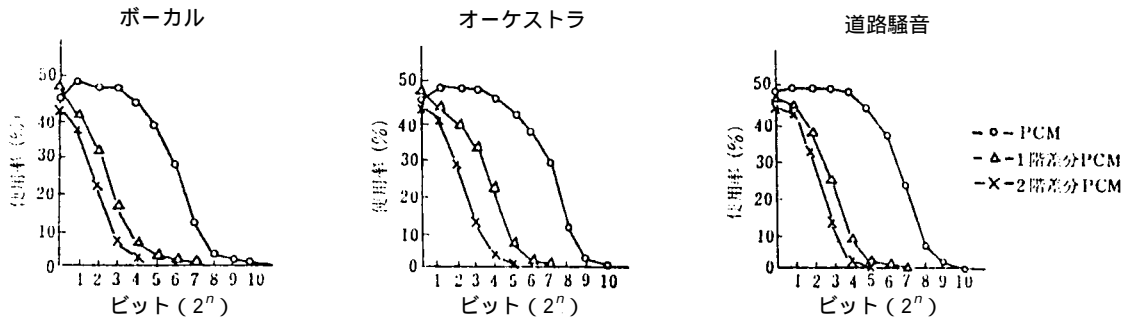


図6 ビット使用率

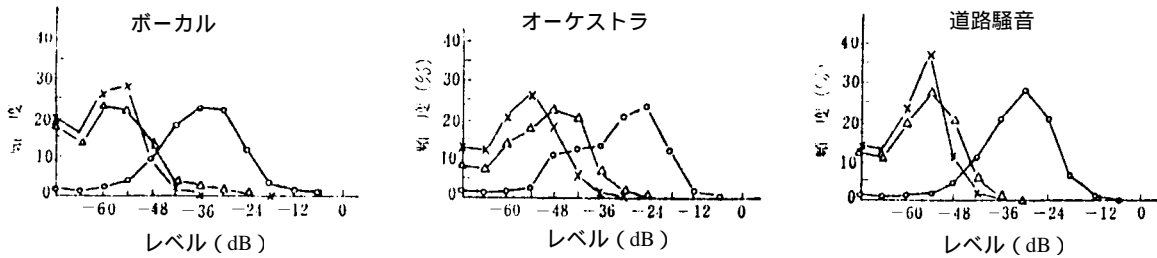


図7 レベル分布

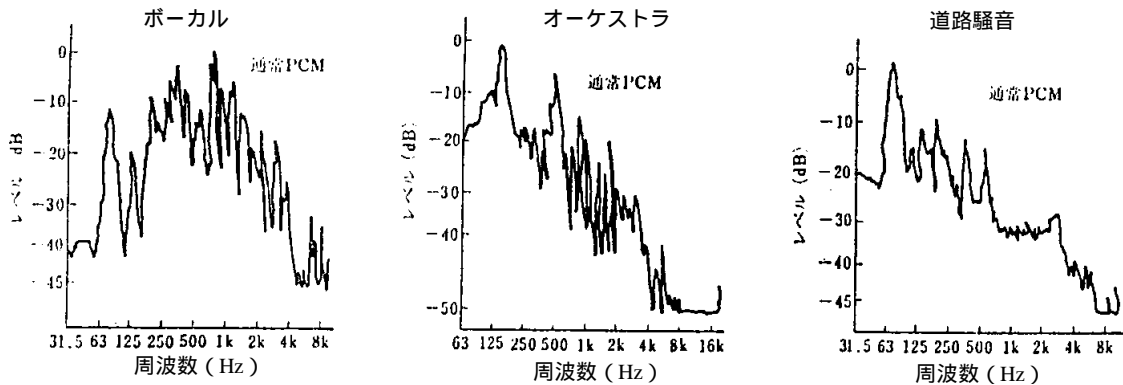


図8 音楽のスペクトル分布

波を付加することがある。BBC では標準化周波数の 1/2 の周波数で量子化ステップの半分の振幅の矩形波と元来の量子化雑音の 4dB の白色雑音を付加することで、S/N 比を 1.5dB 犠牲にして、聴感上 “granular” ひずみを除去している⁽²⁾。

ところで、信号のレベルは大幅に変化するので、振幅の小さい信号に対しては量子化ステップを小さく、振幅の大きい信号に対しては量子化ステップを大きくとれば、小振幅的の SN 比が改善できる。このような量子化を非線形量子化とよぶ。

アナログの瞬時動作の圧伸器を用いても同様の効果を得ることができる。つまり入力標本値のレベル圧縮を行い、復調に際して伸張を行うわけである。いずれにしても、量子化ステップあるいは圧伸関数は、対象とする信号の統計的性質に適したものを、聴感との兼ね合いを含めて慎重に決定する必要がある。

2. 伝送路の高度利用

各種の楽音を標準化周波数 48kHz、量子化ビット数 12、折り返し 2 進符号の PCM 伝送系に与えたときの各ビットの波形を写真 1 に示す。また各ビットの “1” の出現確立を図 6 に、6dB ステップに分けたレベル分布を図 7 に示す。この系の出力をフーリエ変換して得られたスペクトラムのピークを図 8 に示す。

これらの結果から、ビットの使用状況、レベル分布には偏りがあり、スペクトル分布も高域は急激に低下していることがわかる。

一様量子化の通常 PCM では、SN 比は入力振幅に比例し、ダイナミックレンジは直流から限界周波数（標準化周波数の 1/2）まで変わらない。スペクトル分布に着目すれば少なくとも楽音や音声を伝送するうえからは高域までダイナミックレンジが平坦である必要はないし、レベル分布からは、出現頻度の多い中レベル時に SN 比を確保すれば良いとも考えられる。これらの点に着目した伝送路の高度利用、すなわち伝送路の節約が考えられるわけである。

一方、ビット使用率に着目すると、低位のビットはともかく、上位のビットでは “1” の出現確率は非常に小さくなっている。ところで各ビットの 0, 1 が独立に $p, q=1-p$ の確立で出現すると、平均情報量 I は、

$$I = -(p \log_2 p + q \log_2 q) \text{ (ビット)} \quad \dots\dots\dots (5)$$

と表せる。この平均情報量が誤りなく情報を伝送する最低のビットレートを与え、 $p=q=0.5$ のとき $I=1$ ビットとなる。ところで図 6 の測定結果は、0, 1 が独立であると仮定しても、高位のビットに偏りがあるため、全ビット（12 ビットの）総情報量は 8 ~ 9 ビット程度で

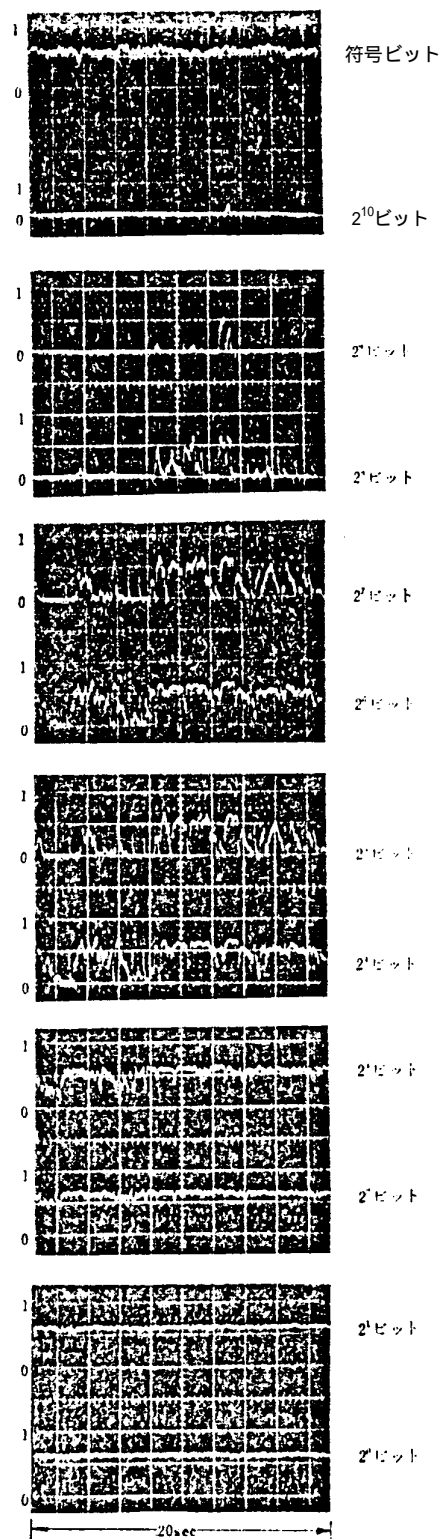
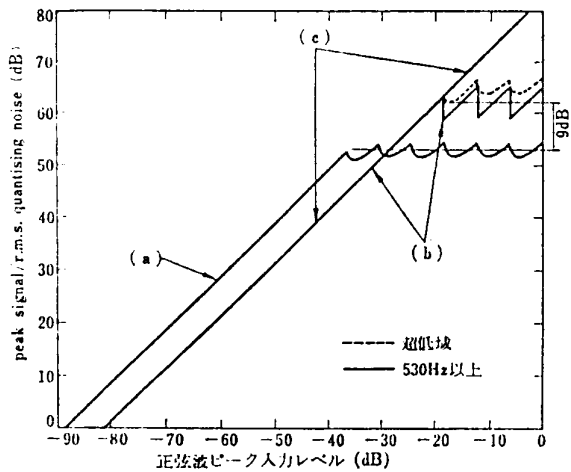


写真 1 各ビットの出力波形（通常 PCM，時定数 20msec の CR 負荷にて撮影）



(a) 瞬時圧伸7セグメント10ビット伝送, (b) Near-instantaneous 4-segment discontinuous law 10.07ビット伝送, (c) 一様符号化13ビット伝送
 図9 BBCのNear-instantaneous 4-segment discontinuous lawによるSN比

ある。その上0, 1は独立ではなくかなり強い相関があることが知られているので平均情報量はさらに減少し4~5ビットである。従って平均情報量からも楽音の伝送については大幅なビットレートの節減の可能性がある。

2.1 非線形量子化

先にも述べたとおり、非線形量子化により、一様量子化に比較して、大振幅時のSN比は悪化するが小振幅時のSN比を改善することが可能である。従って大振幅時のSN比を犠牲にすれば、量子化ビット数を節約することが出来る。

圧伸関数の非線形はデジタル段で与える場合には4~15段階の量子化ステップを使った折れ線を、アナログ段で与える場合には非線形回路素子を用いる。電電公社のPCM-24方式では7ビットで瞬時対数型の圧伸関数を用いている。

最近BBCから、瞬時圧伸ではなく、30標本のレベルを測定して量子化ステップを決定する“Near-instantaneous 4-segment discontinuous law”が報告されている(図9)。それによると10.07ビット相当で最大振幅の-20dB以下では13ビット一様量子化と同等の特性を得ている。

2.2 予測符号化

図8のスペクトラムは楽音の時間相関が強いことを示している。予測符号化は過去の標本値から標本値の予測を行い、真の標本値との差を符号化するものである。従って標本値間の相関が強いほど差信号の振幅が減少し、符号化ビット数の節減が可能である。標本の X_0 予測値

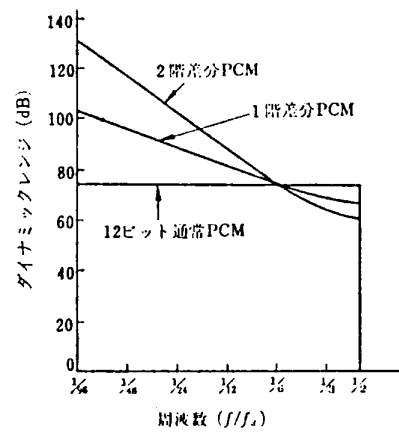


図10 差分PCMのダイナミックレンジ

\hat{X}_0 は,

$$\hat{X}_0 = a_1 X_{-1} + a_2 X_{-2} + \dots + a_n X_{-n} \quad \dots \dots \dots (6)$$

と過去の標本値の一次結合として表される。予測誤差を最小とする a_1, \dots, a_n は標本値間の自己共分散から求められる。実用上からは $a_1 = 0.8 \sim 1$ とし他を0とした一項のみによる予測 $\hat{X}_0 = a_1 X_{-1}$ が広く用いられ、差分PCM (Differential PCM, DPCM) とよばれる。ところで $a_1 = 1$ すなわち隣接標本間の差をとることにより振幅 A_0 , 周波数 h は標本化周波数, $0 \leq h \leq 1/2$ の信号に一度差分をほどこすと振幅 A'_0 は

$$A'_0 = 2A_0 \sin h\pi$$

となる。従って差分PCMにより図8に示すように標本化周波数の1/6を境に、高域でダイナミックレンジは減少し、低域では大幅に増加する。このダイナミックレンジの限界と図8のスペクトル分布から和音に対しては1~2度差分の操作を行うのが有利と考えられる。図6, 図7にデジタル段で構成した差分PCM装置で各種の楽音に差分を1度~2度ほどこしたときの各ビットの“1”出現確率, レベル分布を示すが、差分PCMにより何ら信号に影響をおよぼすことなく3~4ビット節約できることがわかる。

2.3 直交変換

図8のフーリエ変換して得られたスペクトラムは高域の分布の少ないことを示している。従って直交変換を行い系統的にレベルの低い帯域のビット数を減らすことによってビットレートの低減が可能である。直交変換としては Karhunen - Loe've 変換, Fourier 変換, Hadamard 変換等があるが、装置構成上からは, Hadamard 変換あるいは Hadamard-Fourier 変換が有利と考えられる。

2.4 エントロピー符号化

先にも述べたように楽音の平均情報量は4~5ビットである。そこで標本値のとり振幅レベルのうち出現頻度

の大きいレベルに短い符号，小さいものには長い符号を与えるエントロピー符号化を行うことによって，ビットレートを平均情報量と与える限界，Rate distortion boundary に近づけることが可能である。

不等長符号にはシャノンの符号，ハフマン符号等があるが，楽音に対しハフマン符号を適用し，2～3秒のレジスタと組み合わせることにより，5ビットでほぼ完全な伝送が行える⁽³⁾。

3. 各種のデジタル信号処理装置

PCM 信号を扱う際，伝送の過程で，AD 変換，DA 変換をくり返すと SN 比の悪化を招くので，デジタル段階で信号の処理を行った方が有利な場合が多い。また，デジタル処理することによりアナログ系では実現不可能であったような機能の実現もいくつか考えられる。当然，アナログ系で担わせた方が有利な機能も多数ある。

3.1 ミクシング

デジタル信号のミクシングはデジタルの全加算器

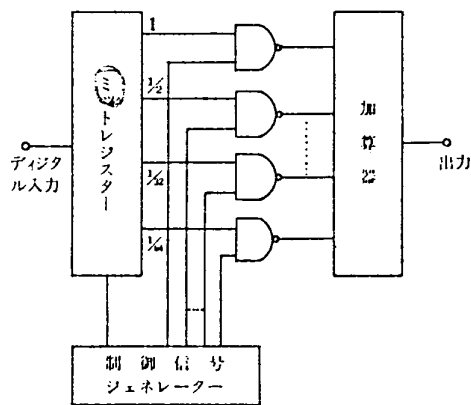


図11 フェーダーのブロック図

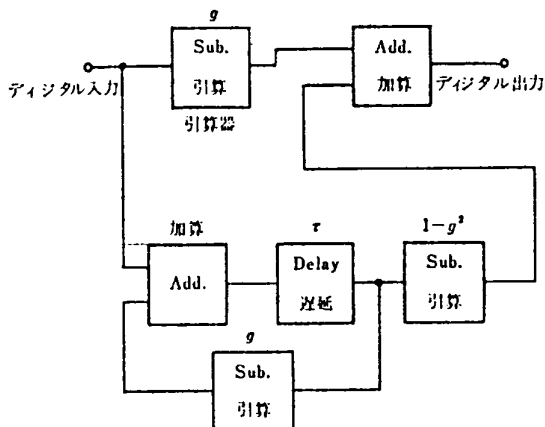


図12 デジタル残響付加装置のブロック図

を用いて簡単に実現できる。この場合，加算器のビット数は信号系 + 1 とする必要がある。

3.2 フェーダー

放送，録音においてもフェーダーは重要な役目を果たすが，特に心理評価の実験等では全てのサンプル自体の長さはもちろんフェーディングのタイミング，傾斜が等しいことが要求される。図11にこうした目的で作られた12ビットのフェーダーのブロック図を示す。シフトレジスタと加減算器，制御回路で構成され，1～1/64の重みをもつ六つのレジスタを加減算し，それを4回繰り返して1dB ± 20%のステップを得ている。フェードイン，アウトの傾斜，タイミングは各々任意に設定できる。レジスタや加算器の数を増やせばステップ幅を小さく精度をより上げることが可能であるが，実用上は十分である。

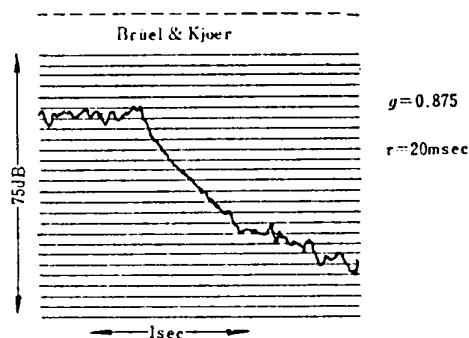


図13 500Hzバンドノイズの残響波形

3.3 遅延装置

遅延装置は，要求が多く，アナログでもテープ式のもの実用に供されているが，可動部分をもち性能上も問題の多い機能である。デジタルではシフトレジスタを用いることにより信号劣化なく任意の遅延を得ることができる。シフトレジスタの価格も大分前に10円/ビットを割り，海外ではA - D，D - A変換器を内蔵したデジタル式遅延装置が数多く市販されている。

3.4 残響付加装置

遅延装置と加算器を組み合わせることにより，図12に示すようなシュレツダの提唱した振幅周波数特性のないいわゆる“カラレス”残響付加装置を忠実に構成することができる。デジタル式では帰還率 g をいくらでも1に近づけることができるので，“残響時間”は自由に選択できる。しかし，特有な“残響音”が作られ，楽器的要素は認められるが，残響付加装置としては疑問が残る。図13に遅延時間 $\tau = 20\text{msec}$ ，帰還率 $g = 0.875$ としたときの500Hzバンドノイズの残響形を示す。

3.5 録音機

アナログのテープレコーダは，ダイナミックレンジ，

ジッター，ドロップアウト，転写，テープによるバイアス，イコライザー特性の相違等問題が多く伝送系の隘路となっている。そこで記録系のデジタル化が考えられる。

デジタル信号を記録するという観点からみれば，コンピュータの磁気テープ記憶装置（MT）や磁気ディスクもPCM記録装置であるが，純然たる広帯域音響信号の記録を目的として作られたのは，1968年NHK技研で開発された簡易型VTRを用いた標準化周波数30kHz，13ビットのPCMレコーダが最初であろう⁽⁴⁾。

その後，1972年日本コロムビアで放送用ローバンドTRと使った。標準化周波数47.25kHz（テレビの垂直同期信号の3倍）13ビットの8チャンネルPCMレコーダが実用化され，これをマスターレコーダとしたレコーダが数多く作られている^{(5)・(6)}。

一方，BBC，カリフォルニア大学，早稲田大学，ソニー，日立等から主として並列記録形のPCMレコーダの実験あるいは実用化が報告されている。ここでは詳細についてはふれないが文献を参照されたい。

3.6 周波数特性可変装置

6dB/octを単位とした振幅周波数特性は，先にふれた微分，および積分装置で簡単に実現できる。従って加減算器と組み合わせることによって，エンファシス特性等を合成することは可能であるが，いわゆるグラフィックイコライザーは直交変換を行う必要がある。

直交変換としては，先にふれたように各種考えられるが，イコライザーとしてはHadamard-Fourier変換が適当であろう。標本値をフーリエ変換して，周波数領域で，振幅，場合によっては位相の値を変化させ逆変換するわけであるが，直交変換する標本の数が，周波数分解能と低域の限界を決定する。50Hz以上（200msec間隔）を直交変換するには400次の直交変換系が必要であり，その実現は容易ではない。

直交変換は先に述べた伝送路の節約や周波数領域で，

振幅や位相を変化させるだけでなく周波数軸の圧縮，伸張が可能なので，ヘリウム音声の復元や特殊な効果音の製作に応用することができる。

む す び

以上，音響信号のデジタル化，PCM伝送デジタル処理について述べた。半導体，集積回路およびその周辺技術の進歩が，従来とうてい実現不可能と思われたような機能をこともなげに実現させて来ている。音響の分野にもますますデジタル技術がとり入れられ，現在すでに普及しつつある遅延装置のように個別の機能のデジタル化も行われることと思われるが，AD，DA変換を繰り返したり標準化周波数やビット数の異なったデジタル機器が同じ系にいくつも入れれば新たな障害が発生する恐れがある。デジタル技術の導入にあたっては，まえがきにもふれたように，デジタル処理，アナログ処理各々に適した機能を見きわめ，各々の特長を十分活用できるように常にシステムバランスを考慮することが望まれる。

【参考文献】

- (1) SHORTER, D.E.L and CHEW, J.R. 1972. Application of pulse-code modulation to sound-signal distribution in a broadcasting network. Proc. IEE, 1972. 119, 10, pp.1442-1448.
- (2) CROLL, M.G. 1970. Pulse-code modulation for high-quality sound signal distribution: quantising distortion at very low signal levels. BBC Research Department Report No. 1970/18.
- (3) 山崎芳男・伊藤毅:広帯域音響信号の高忠実度PCM記録および伝送について，テレビジョン学会，1975年3月。
- (4) 林謙二:PCM磁気録音機，磁気記録研究会資料，MR 69 22.
- (5) 林・宮下・大新田・轟・穴沢:ディスクレコードマスター用PCM録音装置について，音響学会講演論文集，昭和47年10月，p.159
- (6) 穴沢健明:パルス符号変調録音装置による録音，放送技術，1974年10月号。