

HEINRICH·HERTZ·INSTITUT FÜR SCHWINGUNGSFORSCHUNG  
BERLIN·CHARLOTTENBURG

# Technischer Bericht Nr. 150

Simulation von Differenz-PCM-Systemen  
zur Sprachcodierung

von

Dr.-Ing. Peter Noll / Dipl.-Ing. Rolf Block

Berlin

1 9 7 2

Simulation von Differenz-PCM-Systemen zur Sprachcodierung

Zusammenfassung:

Die Differenz-Pulscodemodulation erlaubt eine im Vergleich zu einer Pulscodemodulation genauere digitale Darstellung analoger Signale, weil bei dieser Codierung die dem Signal innewohnende Redundanz reduziert bzw. beseitigt werden kann. Dazu wird aus vorhergehenden Abtastwerten des zu codierenden Signals ein Schätzwert für den nächsten Wert berechnet und nur die Differenz zum wirklich eingetretenen Wert codiert und übertragen.

Die Brauchbarkeit des DPCM-Verfahrens zur Codierung von Sprachsignalen wurde mit Hilfe von Rechnersimulationen untersucht. Die Ergebnisse werden in der vorliegenden Arbeit dargestellt und mit theoretischen Ergebnissen verglichen.

HEINRICH-HERTZ-INSTITUT

Die Bearbeiter

*P. Noll*

(Dr.-Ing. P. Noll)



*R. Block*

(Dipl.-Ing. R. Block)

Der Abteilungsleiter

*E. R. Berger*

(Prof.Dr.-Ing.E.R.Berger)

Der Institutsdirektor

*F. W. Gundlach*

(Prof.Dr.-Ing.F.W.Gundlach)

Berlin-Charlottenburg, den 22. März 1972

## Inhaltsverzeichnis

	Seite
1. <u>Einleitung</u>	1
2. <u>Analyse der Differenz-Pulscodemodulation (DPCM)</u>	3
2.1 Rauschleistung eines PCM-Systems	3
2.2 Rauschleistung eines DPCM-Systems	5
2.3 Vergleich der Rauschleistungen	6
3. <u>Der Entwurf des Prädiktors</u>	7
3.1 Die Aufgabe des Prädiktors	7
3.2 Vereinfachende Annahmen für die DPCM-Sprachcodierung	8
3.3 Bestimmung des optimalen $\underline{h}$ -Vektors	9
3.4 Fehlerrückkopplung	14
3.5 Abnahme des Gewinns bei nichtoptimalem $\underline{h}$ -Vektor	15
3.5.1 Änderung der statistischen Kenndaten des Eingangssignals	15
3.5.2 Änderung des $\underline{h}$ -Vektors	17
3.6 Polynomprädiktion	17
4. <u>Spektrale Darstellung der Differenz-Pulscodemodulation</u>	19
5. <u>Analyse der verwendeten Sprachsignale</u>	22
5.1 Aufnahme der Sprachsignale	22
5.2 Bestimmung der Autokorrelationsfunktion	23
5.3 Bestimmung der Amplitudendichteverteilung	24
6. <u>DPCM-Sprachcodierung</u>	25
6.1 Bisher bekannt gewordene Ergebnisse	25
6.2 Motivation der eigenen Untersuchungen	26
6.3 Simulationsergebnisse	27
6.3.1 Signal-Rausch-Abstände	27
6.3.2 DPCM-Gewinne	33
6.3.3 Weitere Untersuchungen	37
7. <u>Beschreibung des Simulations-Systems</u>	37
8. <u>Literaturverzeichnis</u>	42

## 1. Einleitung

Sollen Sprachsignale in Digitalspeichern gespeichert oder über digitale Kanäle übertragen werden, so müssen die Abtastwerte der Signale in Binärzahlen umgewandelt werden. In den bekannten PCM-Systemen wird jeder Wert mittels einer Quantisierung in einen von  $2^m$  möglichen Werten transformiert und durch eine  $m$ -stellige Binärziffer dargestellt. Das durch die Quantisierung entstehende Quantisierungsrauschen bleibt unter einer zulässigen bzw. wahrnehmbaren Grenze, wenn jeder Abtastwert durch 6 bis maximal 8 Binärstellen dargestellt wird. Eine Verringerung der dabei entstehenden Datenmenge ist oft erwünscht, um den vorgegebenen Speicherplatz bzw. das vorgegebene digitale Übertragungssystem möglichst wirtschaftlich auszunutzen. Da Sprachsignale Redundanzanteile enthalten, ist eine Reduktion der Datenmenge möglich, ohne daß ein vorzugebener, eventuell über der Frequenz bewerteter Fehler zwischen dem Originalsignal und dem rekonstruierten Signal überschritten wird.

Die Aufgaben der Digitalisierung und der Nachrichtenreduktion werden unter dem Begriff der Quellencodierung zusammengefaßt; der Quellencodierer muß die Nachricht der Quelle von der Quellirrelevanz  $H_I$  und von der Redundanz  $R_Q$  befreien /1,2/. Die Irrelevanz  $H_I$  ist der Anteil des Entscheidungsgehaltes  $H_0$  der Quelle, der für den Empfänger entweder nicht interessant oder aber in Kauf zu nehmen ist; die (genügend feine) Quantisierung und die Bandbegrenzung eines Sprachsignals sind Beispiele für eine solche Irrelevanzreduktion. Die Redundanz  $R_Q$  ist der Anteil des Entscheidungsgehaltes  $H_0$  der Quelle, der im Quellencodierer unter Ausnutzung bekannter Eigenschaften des zu codierenden Signals eliminiert und im Quellendecodierer wieder hinzugesetzt werden kann. In der Praxis sind die Irrelevanz- und Redundanzreduktion nicht unabhängig voneinander durchführbar.

Der Entscheidungsgehalt  $H_o = H_I + H_R + R_Q$  wird von dem Quellencodierer im günstigsten Fall um die Anteile  $H_I$  und  $R_Q$  auf die relevante Entropie  $H_R$  reduziert.

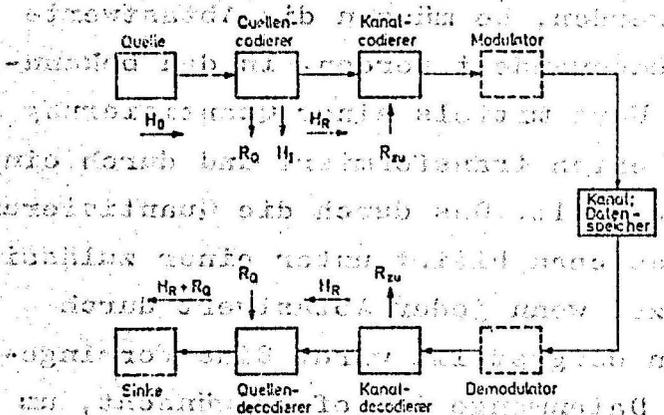


Bild 1 - Quell- und Kanalcodierung

Der binäre Nachrichtenfluß muß anschließend durch eine Kanalcodierung gegen Störungen auf der Übertragungsstrecke bzw. gegen Les- und Schreibfehler bei einer Abspeicherung des Binärsignals gesichert werden. Zur Fehlererkennung und Fehlerkorrektur wird dem binären Signal daher im Kanalcodierer ein Redundanzanteil  $R_{zu}$  zugesetzt /3/.

Die Redundanz eines Sprachsignals besteht aus zwei Anteilen:

- a) Die Amplituden der Abtastwerte der Sprachsignale treten mit unterschiedlicher Wahrscheinlichkeit auf; die Amplitudendichteverteilung kann analytisch angenähert durch eine Laplaceverteilung oder durch eine Gammaverteilung beschrieben werden /4/ (vgl. auch Abschnitt 5.3). Auch bei einer Kompression des Signals treten die Ausgangswerte des Quantisierers i.a. mit unterschiedlicher Wahrscheinlichkeit auf. Die in der ungleichen Verteilung begründete Redundanz kann durch die Huffman- oder durch eine Shannon-Fano-Codierung beseitigt werden /1/, so daß im Prinzip eine Übertragungsrate erreichbar ist, die der Entropie am Quantisiererausgang entspricht. Da die Codeworte bei einer solchen Codierung ungleiche Längen besitzen, muß ein Pufferspeicher verwendet werden, durch den der praktische Aufbau sehr aufwendig wird.

- b) Abtastwerte der Sprachsignale sind von vorhergehenden Abtastwerten statistisch abhängig. Diese Abhängigkeit läßt eine gewisse Vorhersagbarkeit eines neuen Wertes aus bereits bekannten Abtastwerten zu; diese Vorhersagbarkeit läßt sich auf einfache Weise zur Redundanzreduktion ausnutzen.

Die vorliegende Arbeit zeigt, wie die zwischen den Abtastwerten von Sprachsignalen bestehende lineare Abhängigkeit mit Hilfe einer Differenz-Pulscodemodulation (DPCM) weitgehend beseitigt werden kann. Der informationstragende Anteil kann dann bei gleicher Übertragungsrage genauer beschrieben werden, so daß sich das Quantisierungsrauschen gegenüber einer Pulscodemodulation (PCM) verringert. Redundanzanteile, die durch Abhängigkeiten höherer statistischer Ordnung bedingt sind, können durch eine nichtlineare Codiereinrichtung erfaßt und beseitigt werden /5/; diese Möglichkeiten wurden bei unseren Untersuchungen nicht berücksichtigt. Auch auf die Möglichkeit, eine weitere Redundanzreduktion durch eine Umcodierung der Binär-Ausgangswerte des DPCM-Systems zu erzielen /6/, wird im folgenden nicht eingegangen.

## 2. Analyse der Differenz-Pulscodemodulation (DPCM)

Die mit DPCM-Systemen gegenüber PCM-Systemen erreichbare Verringerung des Quantisierungsrauschens soll berechnet werden /7,8/. Es werden nur Systeme mit mehrstufigem Quantisierer betrachtet, d.h., die Deltamodulation mit ihrem 1-bit-Quantisierer (Komparator) bleibt unbeachtet.

### 2.1 Rauschleistung eines PCM-Systems

Im folgenden wird Sprache als stationärer Prozeß vorausgesetzt.  $\{s(k)\}$  sei eine zeitdiskrete Musterfolge dieses Prozesses mit dem Mittelwert  $E[s(k)] = 0$  und der Varianz  $S = \sigma^2 = E[s^2(k)]$ . Jeder Wert  $s(k)$  wird quantisiert, der Index der Quantisierungsstufe wird übertragen.

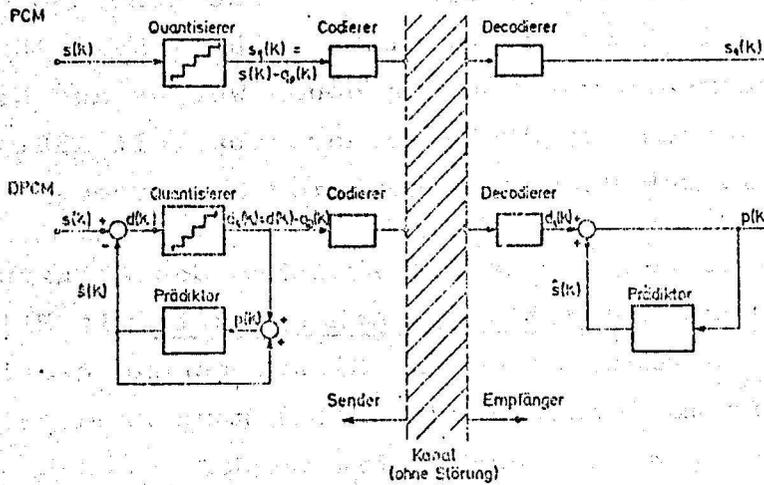


Bild 2  
Vergleich PCM-DPCM

Im Decodierer des Empfängers wird bei fehlerfreier Übertragung der Wert  $s_q(k)$  rekonstruiert, der Gesamtfehler ist also mit dem Quantisierungsfehler identisch:

$$\begin{aligned} q_p(k) &= \varepsilon(k) - s_q(k) \\ &= s(k) - g(s(k)) \end{aligned}$$

$g(\cdot)$  beschreibt die Kennlinie des Quantisierers. Die Varianz des Quantisierungsfehlers ergibt sich zu

$$Q_p = E [ q_p^2(k) ] = S \cdot e_p^2$$

$e_p^2$  ist die normierte, d.h. auf eine Eingangsvarianz 1 bezogene Rauschleistung; sie ist eine Funktion der Quantisierer-Kennlinie  $g(\cdot)$ , der Amplitudendichteverteilung  $p(\cdot)$  und des Aussteuerungsfaktors  $s_c/\sigma$ :

$$e_p^2 = f \{ g(\cdot), p(\cdot), s_c/\sigma \}$$

$s_c$  ist der maximale Amplitudenwert, für den noch keine Begrenzung eintritt (Bild 3).

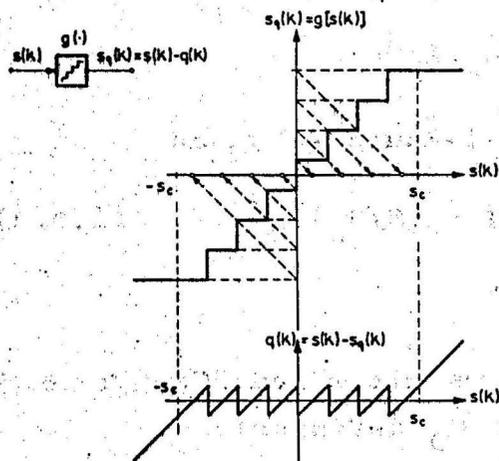


Bild 3  
Quantisierer-Kennlinie und  
Quantisierungsfehler

In PCM-Systemen entsteht daher ein Signal-Rausch-Abstand  
 $S/Q_p = e_p^{-2}$ .

## 2.2 Rauschleistung eines DPCM-Systems

In einem DPCM-System wird von jedem Eingangswert  $s(k)$  ein Schätzwert  $\hat{s}(k)$  subtrahiert (s. Bild 2). Die Differenz  $d(k) = s(k) - \hat{s}(k)$  wird quantisiert ( $d_q(k) = g(d(k))$ ), codiert und übertragen. Im Empfänger wird zu dem ankommenden Wert  $d_q(k)$  ein Schätzwert  $\hat{s}(k)$  addiert, der aus  $N$  vorhergehenden rekonstruierten Werten  $p(j)$ ;  $j = k-1, k-2, \dots, k-N$  ermittelt wurde. Um ein gleichartiges Arbeiten von Sender und Empfänger sicherzustellen, wird auch im Sender die Folge der Rekonstruktionswerte  $\{p(k)\}$  erzeugt und zur Bestimmung des Schätzwertes  $\hat{s}(k)$  herangezogen.

Auch im DPCM-System ist der Gesamtfehler  $s(k) - p(k)$  mit dem Quantisierungsfehler identisch:

$$s(k) - p(k) = s(k) - [d_q(k) + \hat{s}(k)] = q_D(k) \quad (2.2.1)$$

Ist  $S_D = \sigma_D^2 = E[d^2(k)]$  die Eingangsleistung des Quantisierers, so ergibt sich für die Varianz des Quantisierungsfehlers:

$$Q_D = E[q_D^2(k)] = S_D \cdot e_D^2 \quad (2.2.2)$$

$e_D^2$  ist wiederum die auf die Varianz 1 des Quantisierer-Eingangssignals bezogene Rauschleistung.

$$e_D = f \{ g(\cdot), p(\cdot), s_c/\zeta_D \}$$

In DPCM-Systemen entsteht ein Signal-Rausch-Abstand

$$S/Q_D = (S/S_D) \cdot (S_D/Q_D) = (S/S_D) e_D^{-2} \quad (2.2.3)$$

### 2.3 Vergleich der Rauschleistungen

Um die Eigenschaften des DPCM-Systems mit einem PCM-System vergleichen zu können, wird ein Gewinn  $G_D$  definiert:

$$G_D = \left. \frac{Q_p}{Q_D} \right|_{Q_p = \text{const}} = \frac{S}{S_D} \cdot \left. \frac{e_p^2}{e_D^2} \right|_{e_p^2 = \text{const}} \quad (2.3.1)$$

Der mit einem DPCM-System erzielbare Gewinn ist durch das Produkt von zwei Faktoren gegeben:

Der erste Faktor  $S/S_D$  ist abhängig von der Fähigkeit des Prädiktors, aus rekonstruierten Werten einen Schätzwert zu bestimmen, der zu einer geringen Varianz des Differenzsignals führt.

Der zweite Faktor  $e_p^2/e_D^2$  ist der Quotient der normierten Quantisierer-Rauschleistungen des PCM- bzw. des DPCM-Systems. Die Größe  $e_D^2$  kann bei vorgegebener Amplitudendichteverteilung durch eine geeignete Wahl der Quantisierer-Kennlinie näherungsweise /9/, für gaußverteilte Prozesse auch exakt minimisiert werden /10/; es ist jedoch nicht möglich, für die beiden Faktoren  $S/S_D$  bzw.  $e_p^2/e_D^2$  getrennt das jeweilige Maximum anzugeben, da  $S_D$  und  $e_D$  voneinander abhängen: da der Schätzwert  $\hat{s}(k)$  aus rekonstruierten Werken  $p(j)$ ;  $j = k-1, \dots, k-N$  gewonnen wird, die nach Gl.2.2.1 einen Rauschanteil enthalten, ist die Optimierung des Prädiktors von der Rauschleistung  $Q_D = S_D \cdot e_D^2$  abhängig. Andererseits beeinflusst die Dimensionierung des Prädiktors auch die Amplitudendichteverteilung des Differenzsignals  $\{d(k)\}$ , die wiederum für die Größe der normierten Rauschleistung  $e_D^2$  wesentlich ist. Bei

der Codierung von Sprachsignalen kann jedoch eine wesentliche Vereinfachung durch den Ansatz  $e_D^2 \approx e_p^2$  erzielt werden. Im nächsten Abschnitt wird diese Annahme begründet.

### 3. Der Entwurf des DPCM-Prädiktors

#### 3.1 Die Aufgabe des Prädiktors

Der Prädiktor des DPCM-Systems ist ein nichtrekursives Filter, das aus rekonstruierten Werten  $p(j)$ ;  $j=k-1, k-2, \dots, k-N$  einen Schätzwert  $\hat{s}(k)$  als Linearkombination

$$\hat{s}(k) = \sum_{j=1}^N h_j \cdot p(k-j) = \underline{h}^T \cdot \underline{p}(k) \quad (3.1.1)$$

gewinnt.

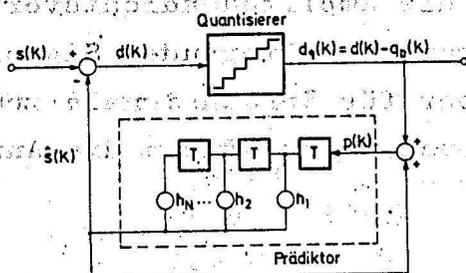


Bild 4

DPCM-System

Aufbau des Prädiktors

$\underline{h}^T = (h_1, h_2, \dots, h_N)$  ist der Vektor der Prädiktorkoeffizienten,  $\underline{p}(k)$  mit den Elementen  $p(j)$ ;  $j=k-1, k-2, \dots, k-N$  ist der Vektor der rekonstruierten Werte, die im Prädiktor gespeichert sind. Die Aufgabe des Prädiktor-Entwurfs besteht darin, einen Vektor  $\underline{h}$  zu finden, der den Gewinn  $G_D$  nach Gl. 2.3.1 zum Maximum macht, wobei die Eigenschaften des PCM-Systems als vorgegeben anzusehen sind.

$$G_D = \frac{S}{S_D} \cdot \frac{e_p^2}{e_D^2} \Big|_{e_p^2 = \text{const}} = \text{max.} \quad (3.1.2)$$

### 3.2 Vereinfachende Annahmen für die DPCM-Sprachcodierung

Wegen der im Abschnitt 2.3 beschriebenen Abhängigkeit zwischen der Varianz des Differenzsignals und der normierten Rauschleistung des DPCM-Quantisierers ist es nicht möglich, eine Berechnung des maximal möglichen Gewinns nach Gl. 2.3.1 und eine Bestimmung des optimalen  $h$ -Vektors vorzunehmen. Bei Sprachsignalen, die im Zuge einer Übertragung über Fernsprechanäle einem Codierer angeboten werden, kann der mittlere Pegel großen Schwankungen von mehr als 30 dB unterliegen. Daher ist der Quantisierer-Kennlinie  $g(\cdot)$  ein logarithmischer Verlauf vorgeschrieben, der zu einem über größere Bereiche der Eingangsleistung konstantem Signal-Rausch-Abstand führt /11,12/, (vgl. auch Abschnitt 6.3.1). Sind aber die Kennlinien des PCM- und des DPCM-Quantisierers gleich, so ergibt sich dann für beide Systeme eine gleiche normierte Rauschleistung, wenn die Amplitudendichteverteilungen der an den Quantisierer-Eingängen anliegenden Signale gleich sind. Diese Bedingung ist aber für Sprachsignale zumindest grob erfüllt /13/; eigene Messungen bestätigen die Annahme:

$$e_D^2 \approx e_p^2$$

Bild 5 zeigt entsprechende Ergebnisse (17sec-Simulationsläufe für einen Sprecher). Die Bestimmung der

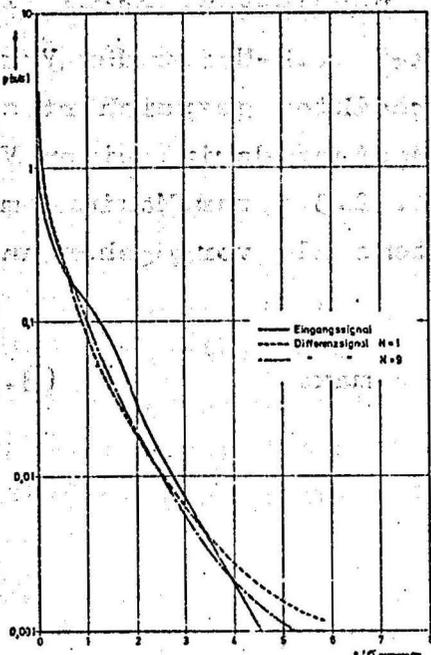


Bild 5  
Amplitudendichteverteilungen  
Sprecher 13 (TP-Sprache)

Amplitudendichteverteilung wird im Abschnitt 5.3 beschrieben.

Damit kann für Sprachsignale ein Gewinn

$$G_S = \frac{S}{S_D} \quad (3.2.1)$$

definiert werden. Es ist daher die Aufgabe des Prädiktor-Entwurfs, denjenigen  $\underline{h}$ -Vektor zu finden, der  $S_D$ , die Varianz des Differenzsignals, zum Minimum macht. Dabei wird nicht als Nebenbedingung gefordert, daß das Differenzsignal optimal quantisiert wird.

### 3.3 Bestimmung des optimalen $\underline{h}$ -Vektors

Bei vorgegebener Varianz des Eingangssignals wird der Gewinn  $G_S$  nach Gl. 3.2.1 ein Maximum, wenn die Varianz des Differenzsignals zum Minimum wird.

Es gilt

$$\begin{aligned} S_D &= E [ d^2(k) ] = E [ \{ s(k) - \hat{s}(k) \}^2 ] \\ &= E [ \{ s(k) - \underline{h}^T \cdot \underline{p}(k) \}^2 ] \\ &= S - 2\underline{h}^T E [ s(k) \cdot \underline{p}(k) ] + \underline{h}^T E [ \underline{p}(k) \cdot \underline{p}^T(k) ] \cdot \underline{h} \end{aligned} \quad (3.3.1)$$

Wie aus Gl. 2.2.1 ersichtlich, gilt für den Vektor  $\underline{p}$  :

$$\underline{p}(k) = \underline{s}(k) - \underline{q}_D(k)$$

$\underline{s}(k)$  = Vektor mit den Elementen  $s(j)$ ;  $j=k-1, k-2, \dots, k-N$

$\underline{q}_D(k)$  = Vektor mit den Elementen  $q_D(j)$ ;  $j=k-1, k-2, \dots, k-N$

#### Erste Vereinfachung

Um die Varianz des Differenzsignals mit Gl. 3.3.1 berechnen zu können, ist eine Kenntnis der Kreuzkorrelation zwischen den Werten des Eingangssignals und den Werten des Quantisierungsfehlers erforderlich. Es kann aber fast immer vorausgesetzt werden, daß diese Prozesse nicht miteinander korreliert sind:

$$E [ s(k) \cdot q_D(l) ] = 0 \quad \forall k, l$$

Damit wird Gl. 3.3.1 zu

$$\begin{aligned}
 S_D &= S - 2\mathbf{h}^T E[\mathbf{s}(k) \cdot \underline{\mathbf{s}}(k)] + \mathbf{h}^T E[\underline{\mathbf{s}}(k) \underline{\mathbf{s}}^T(k) + \mathbf{q}_D(k) \mathbf{q}_D^T(k)] \cdot \mathbf{h} \\
 &= S \left\{ 1 - 2\mathbf{h}^T \underline{\mathbf{r}}_S + \mathbf{h}^T (\underline{\mathbf{R}}_S + \frac{Q_D}{S} \underline{\mathbf{R}}_Q) \mathbf{h} \right\} \\
 &= S \left\{ 1 - 2\mathbf{h}^T \underline{\mathbf{r}}_S + \mathbf{h}^T \underline{\mathbf{R}} \mathbf{h} \right\} \quad (3.3.2)
 \end{aligned}$$

$\underline{\mathbf{r}}_S$  = Vektor der normierten AKF-Werte der Eingangsfolge mit den Elementen  $r_S(j) = E[\mathbf{s}(k) \cdot \mathbf{s}(k-j)]/S = \rho(j)$ ;  $j=1,2,\dots,N$

$\underline{\mathbf{R}}_S = E[\underline{\mathbf{s}}(k) \cdot \underline{\mathbf{s}}^T(k)]/S$  = Kovarianzmatrix der Eingangsfolge mit den normierten Elementen  $E[\mathbf{s}(k-i) \mathbf{s}(k-j)]/S$ ;  $i, j=1,2,\dots,N$

$\underline{\mathbf{R}}_Q = E[\mathbf{q}_D(k) \cdot \mathbf{q}_D^T(k)]/Q_D$  = Kovarianzmatrix der Folge der Quantisierungsfehler mit den normierten Elementen  $E[\mathbf{q}_D(k-i) \cdot \mathbf{q}_D(k-j)]/Q_D$ ;  $i, j=1,2,\dots,N$

$$\underline{\mathbf{R}} = \underline{\mathbf{R}}_S + (Q_D/S) \cdot \underline{\mathbf{R}}_Q = \underline{\mathbf{R}}_S + (e_D^2/G_S) \cdot \underline{\mathbf{R}}_Q \quad (3.3.3)$$

Aus Gl. 3.3.2 folgt mit den Gl. 2.2.2, 2.2.3 und 3.2.1

$$G_S = S/S_D = (1 - 2\mathbf{h}^T \underline{\mathbf{r}}_S + \mathbf{h}^T \underline{\mathbf{R}} \mathbf{h})^{-1} \quad (3.3.4a)$$

$$\boxed{G_S = \frac{1 - e_D^2 \mathbf{h}^T \underline{\mathbf{R}}_Q \mathbf{h}}{1 - 2\mathbf{h}^T \underline{\mathbf{r}}_S + \mathbf{h}^T \underline{\mathbf{R}}_S \mathbf{h}}} \quad (3.3.4b)$$

und der Signal-Rausch-Abstand

$$S/Q_D = (S/S_D) e_D^{-2} = G_S / e_D^2 \quad (3.3.5)$$

Aus Gl. 3.3.4a kann im Prinzip der optimale Koeffizientenvektor  $\underline{\mathbf{h}}_0$  bestimmt werden, der den Gewinn zum Maximum macht, wenn vorausgesetzt wird, daß die normierte Rauschleistung  $e_D^2$  nur wenig von der Größe der Vektoren  $\underline{\mathbf{h}}$  beeinflusst wird. Die Optimierung des Gewinns bezüglich des Vektors  $\underline{\mathbf{h}}$  führt jedoch auf

eine Matrizengleichung, die eine explizite Angabe des  $\underline{h}_0$ -Vektors nicht zuläßt. Bei den Simulationsläufen wurde daher ein suboptimaler Vektor  $\underline{h}_*$  bestimmt, der nur von den statistischen Kenndaten der Eingangsfolge abhängt:

### Zweite Vereinfachung

Es wird derjenige Vektor  $\underline{h}_*$  bestimmt, der die Varianz des Differenzsignals zum Minimum macht (und damit den Gewinn maximiert), wenn der Einfluß des Quantisierers vernachlässigbar ist ( $Q_D \rightarrow 0$ ). Aus den Gleichungen 3.3.2 und 3.3.3 folgt

$$S_D = S(1 - 2\underline{h}^T \underline{r}_S + \underline{h}^T \underline{R}_S \underline{h})$$

$S_D = S_D(\underline{h})$  muß nach den einzelnen Komponenten  $h_j$ ;  $j=1,2,\dots,N$  abgeleitet und jeweils zu Null gesetzt werden:

$$\nabla S_D(\underline{h}) = 0$$

ergibt das Gleichungssystem

$$\underline{r}_S = \underline{R}_S \cdot \underline{h}_* \quad (3.3.6)$$

Die Kovarianzmatrix  $\underline{R}_S$  ist positiv-definit, daher wird

$$\underline{h}_* = \underline{R}_S^{-1} \underline{r}_S \quad (3.3.7)$$

$\underline{h}_*$  ist der optimale Koeffizientenvektor, wenn die Prädiktion mit vorhergehenden Eingangswerten vorgenommen wird. Werden die rekonstruierten Werte  $p(j)$ ;  $j=k-1, k-2, \dots, k-N$  zur Prädiktion herangezogen, so bleiben die Abweichungen zwischen  $\underline{h}_*$  und dem optimalen Vektor  $\underline{h}_0$  (der aus Gl. 3.3.4b zu bestimmen wäre) klein, wenn der Signal-Rausch-Abstand genügend groß ist ( $Q_D \ll S$ ). Wird der gemäß Gl. 3.3.7 aus den Autokorrelationswerten der Eingangsfolge  $\{s(k)\}$  bestimmte Vektor  $\underline{h}_*$  zur Schätzung herangezogen, so ergibt sich aus Gl. 3.3.4 der suboptimale Gewinn

$$G_{Su} = \left(1 - \frac{h^T R_S}{S} + \frac{Q_D}{S} \frac{h^T R_Q h}{h^T R_S h}\right)^{-1}$$

$$G_{Su} = \frac{1 - e_D^2 \frac{h^T R_Q h}{h^T R_S h}}{1 - \frac{h^T R_S}{S} \frac{h^T R_Q h}{h^T R_S h}} \quad (3.3.8)$$

und der suboptimale Signal-Rausch-Abstand

$$S/Q_{Du} = \frac{G_{Su}}{e_D^2} = \frac{1 - e_D^2 \frac{h^T R_Q h}{h^T R_S h}}{e_D^2 (1 - \frac{h^T R_S}{S} \frac{h^T R_Q h}{h^T R_S h})} \quad (3.3.9)$$

Wird zur Vorhersage nur der letzte Abtastwert herangezogen (N=1), so vereinfachen sich die beiden letzten Gleichungen zu

$$G_{Su} = \frac{1 - e_D^2 h_1^2}{1 - h_1^2}$$

und zu

$$S/Q_{Du} = \frac{1 - e_D^2 h_1^2}{(1 - h_1^2) e_D^2}$$

Die letzte Gleichung wurde bereits von O'Neal angegeben /14/.

Ist der Einfluß des Quantisierers ganz vernachlässigbar, so entsteht der maximal mögliche Gewinn bzw. Signal-Rausch-Abstand

$$G_{S^*} = G_{Su} \Big|_{Q_D \ll S} = (1 - \frac{h^T R_S h}{S})^{-1} = (1 - \frac{h^T R_S}{S})^{-1} \quad (3.3.10)$$

$$S/Q_{D^*} = G_{S^*} / e_D^2 \quad (3.3.11)$$

### Dritte Vereinfachung

Sind die bei der Quantisierung entstehenden Quantisierungsfehler nicht miteinander korreliert,

$$E [ q_D(k) \cdot q_D(l) ] = q_D \cdot \delta_{kl} ,$$

so wird die Kovarianzmatrix  $\underline{R}_Q$  zur Einheitsmatrix  $\underline{I}$ . Das Gleichungssystem 3.3.6 kann dann in der Form

$$r_{Sm}(j) = \sum_{i=1}^N h_{io} \cdot r_{Sm}(j-i) \quad j=1,2,\dots,N \quad (3.3.12)$$

geschrieben werden, wobei die Koeffizienten  $h_{io}$ ;  $i=1,2,\dots,N$  Elemente des Vektors  $\underline{h}_o$  sind. Die modifizierten AKF-Werte  $r_{Sm}(j)$  sind durch

$$\begin{aligned} r_{Sm}(0) &= 1 & (3.3.13) \\ r_{Sm}(j) &= \frac{r_S(j)}{1+Q_D/S} \quad j \neq 0 \end{aligned}$$

gegeben; diese Darstellung wurde zuerst von McDonald verwendet /13/. Da die Schätzwerte  $\hat{s}(k)$  aus rekonstruierten, d.h. mit einem Quantisierungsfehler behafteten Werten  $p(j)$ ;  $j=k-1,\dots,k-N$  bestimmt werden müssen, kann die Korrelation der Eingangsfolge  $\{s(k)\}$  nicht voll zur Schätzung ausgenutzt werden. Es ergibt sich daher als äquivalente Darstellung gemäß den Gleichungen 3.3.12 und 3.3.13 eine optimale Schätzung aus einer Folge mit verringerten Autokorrelationswerten. Bild 6 zeigt, wie sich die modifizierte Autokorrelationsfunktion eines Sprachsignals in Abhängigkeit von dem Signal-Rausch-Abstand ändert (AKF-Berechnung: siehe Abschnitt 5.2).

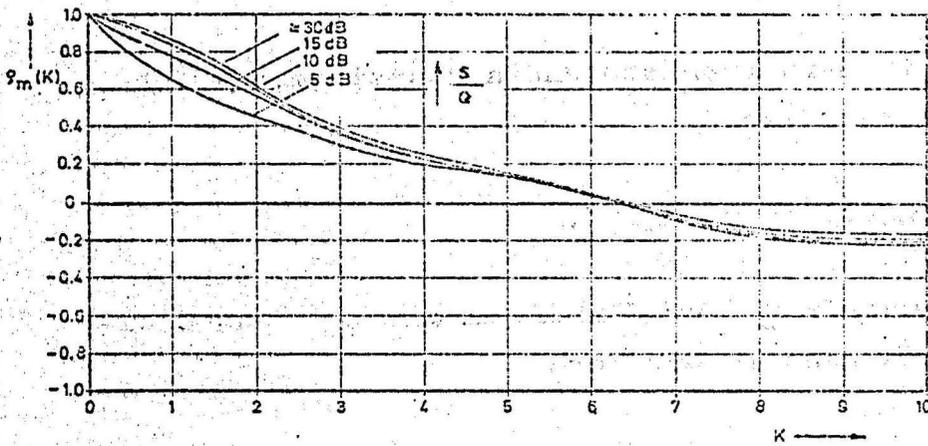


Bild 6  
 modifizierte Autokorrelationsfunktionen nach  
 Gl. (3.3.13) : 125  $\mu$ s-  
 Raster  
 Sprecher 13 (TP-Sprache)

### 3.4 Fehlerrückkopplung

Nach Gl. 3.3.2 ergibt sich die Varianz des Differenzsignals zu

$$S_D = S(1 - 2\underline{h}^T \underline{r}_S + \underline{h}^T \underline{R} \underline{h})$$

Ist der Einfluß des Quantisierers vernachlässigbar ( $Q_D \ll S$ ;  $\underline{R} = \underline{R}_S$ ), so wird  $S_D$  nach Einsetzen des optimalen Vektors  $\underline{h}_* = \underline{R}_S^{-1} \underline{r}_S$  zu

$$S_{D*} = S(1 - \underline{h}_*^T \underline{r}_S)$$

Für die suboptimale Varianz folgt

$$\begin{aligned} S_{Du} &= S(1 - 2\underline{h}_*^T \underline{r}_S + \underline{h}_*^T (\underline{R}_S + (Q_D/S) \underline{R}_Q) \underline{h}_*) \\ &= S(1 - \underline{h}_*^T \underline{r}_S + (Q_D/S) \underline{h}_*^T \underline{R}_Q \underline{h}_*) \\ &= S_{D*} + Q_D \underline{h}_*^T \underline{R}_Q \underline{h}_* \end{aligned} \quad (3.5.1)$$

Durch die Rückkopplung des Quantisierungsfehlers erhöht sich somit die Varianz des Differenzsignals um die quadratische Form  $Q_D \underline{h}_*^T \underline{R}_Q \underline{h}_*$ . Der erreichbare Gewinn wird entsprechend von  $G_{S*}$  auf  $G_{Su}$  verringert.

$$G_{Su} = G_{S*} (1 - e_{D \underline{h}_*^T \underline{R}_Q \underline{h}_*}) \quad (3.5.2)$$

Die Verringerung des maximal möglichen Gewinnes kann einmal so gedeutet werden, daß zur Prädiktion nur rauschbehaftete Werte  $p(j)$ ;  $j = k-1, k-2, \dots, k-N$  zur Verfügung stehen; zum andern kann die Verringerung dadurch erklärt werden, daß sich die Varianz des Differenzsignals um die Varianz des rückgekoppelten und gefilterten Quantisierungsfehlers erhöht hat. Beide Anschauungen führen natürlich zu dem gleichen Ergebnis (Gl. 3.3.8 bzw. Gl. 3.5.2).

### 3.5 Abnahme des Gewinnes bei nichtoptimalem $\underline{h}$ -Vektor:

#### 3.5.1 Änderung der statistischen Kenndaten des Eingangssignals

Ein DPCM-Codierer muß Sprachsignale verschiedener Sprecher mit unterschiedlichen Autokorrelationsfunktionen verarbeiten; bei einem nichtadaptiven Codierer liegt der  $\underline{h}$ -Vektor fest, er wird gemäß dem Mittelwert der statistischen Kenndaten verschiedener Sprecher festgelegt. Wenn die linearen Abhängigkeiten des Eingangssignals durch  $\underline{r}_S$  bzw.  $\underline{R}_S$  beschrieben werden können, dann würden sich der (sub-)optimale Vektor  $\underline{h}_*$  und der maximale Gewinn  $G_*$  gemäß Gl. 3.3.7 bzw. Gl. 3.3.10 ergeben. Da der Prädiktor auf die Koeffizienten  $\underline{h}$  eingestellt ist, ergibt sich mit Gl. 3.3.4b ein reduzierter Gewinn

$$G' = \frac{1 - e_D^2 \underline{h}^T \underline{R}_Q \underline{h}}{1 - 2 \underline{h}^T \underline{r}_S' + \underline{h}^T \underline{R}_S' \underline{h}} \quad (3.5.1)$$

Wir führen einen Fehlervektor  $\underline{f}$  ein, um den sich der eingestellte Koeffizientensatz von dem optimalen Koeffizientensatz unterscheidet.

$$\underline{h}_* = \underline{h} + \underline{f} \quad (3.5.2)$$

Der Nenner von Gl. 3.5.1 läßt sich wie folgt vereinfachen.

$$\begin{aligned}
 & 1 - 2\underline{h}^T \underline{r}'_S + \underline{h}^T \underline{R}'_S \underline{h} \\
 &= 1 - 2\underline{h}_*^T \underline{r}'_S + \underline{h}_*^T \underline{R}'_S \underline{h}_* + 2\underline{f}^T \underline{r}'_S - \underline{f}^T \underline{R}'_S \underline{h}_* - \underline{h}_*^T \underline{R}'_S \underline{f} + \underline{f}^T \underline{R}'_S \underline{f} \\
 &= G_*'^{-1} + 2\underline{r}'_S{}^T \underline{f} - 2\underline{h}_*^T \underline{R}'_S \underline{f} + \underline{f}^T \underline{R}'_S \underline{f} \\
 &= G_*'^{-1} + 2(\underline{r}'_S - \underline{R}'_S \underline{h}_*)^T \underline{f} + \underline{f}^T \underline{R}'_S \underline{f} \\
 &= G_*'^{-1} + \underline{f}^T \underline{R}'_S \underline{f}
 \end{aligned}$$

Damit ergibt sich der reduzierte Gewinn

$$G' = \frac{G_*' (1 - e_D^2 \underline{h}^T \underline{R}_Q \underline{h})}{1 + G_*' \underline{f}^T \underline{R}'_S \underline{f}} \quad (3.5.3)$$

Bild 7 zeigt, wie sich für vier Sprecher (AKF dazu: siehe Abschnitt 5.2) der Gewinn verringert, wenn der Koeffizientensatz gemäß dem Mittelwert der Autokorrelationsfunktion der vier Sprecher eingestellt ist.

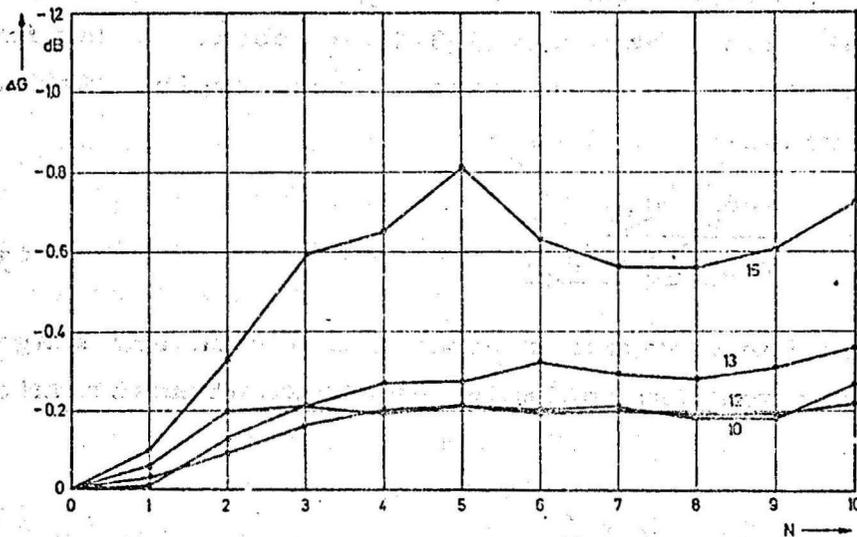


Bild 7 Abnahme des Gewinns durch Einstellung der Prädiktor-koeffizienten gemäß dem Mittelwert der AKF für 4 Sprecher

### 3.5.2 Änderung des $\underline{h}$ -Vektors

Eine Reduktion des Gewinnes tritt auch ein, wenn der  $\underline{h}$ -Vektor bei gleichbleibendem Eingangssignal geändert wird. Eine solche Änderung ergibt sich z.B. bei einer digitalen Realisierung des Prädiktors; es entsteht dann ein Rundungsfehler, wenn für die Darstellung der Prädiktorkoeffizienten nur eine endliche Zahl von bit zur Verfügung stehen /7/. Es entsteht ein neuer Vektor  $\underline{h}$ , der sich von der optimalen Lösung um den Vektor der Rundungsfehler  $\underline{r}$  unterscheidet:

$$\underline{h} = \underline{h}_* + \underline{r} \quad (3.5.4)$$

Der reduzierte Gewinn ergibt sich zu

$$G = \frac{G_* \left( 1 - e_D^2 \underline{h}^T \underline{R}_Q \underline{h} \right)}{1 + G_* \underline{r}^T \underline{R}_S \underline{r}} \quad (3.5.5)$$

### 3.6 Polynomprädiktion

Durch N im Prädiktor gespeicherte Werte  $p(j)$ ;  $j=k-1, k-2, \dots, k-N$  wird ein Polynom  $(N-1)$ ter Ordnung gelegt und eine Extrapolation auf den nächsten Wert hin vorgenommen. Die Newton'sche Interpolation für äquidistante Stützstellen liefert die gesuchten Komponenten  $h_j$  des Koeffizientenvektors  $\underline{h}$ :

$$h_j = (-1)^{j+1} \binom{N}{j} \quad (3.6.1)$$

Die Koeffizienten sind unabhängig von der Statistik des Eingangssignals, jedoch wird der Gewinn weiterhin von den statistischen Kenndaten des Eingangssignals bestimmt. Mit dem Ansatz nach Gl. 3.5.4 kann er aus der Gl. 3.5.5 bestimmt werden.

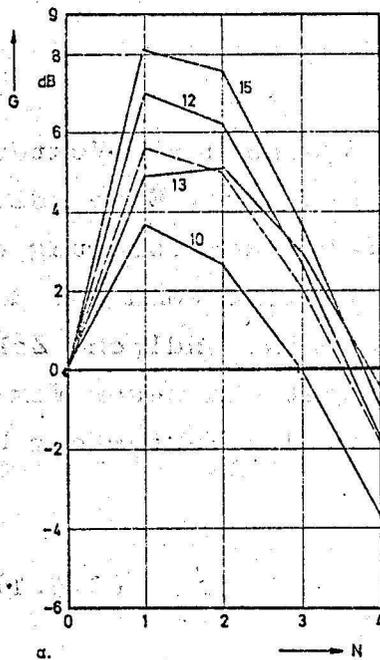


Bild 8 zeigt die Abhängigkeit dieses Gewinnes von der Zahl der Prädiktor-koeffizienten für vier Autokorrelations-funktionen, die im Bild 11 dargestellt sind (Sprachproben der Sprecher 10,12, 13,15). Die gestrichelte Kurve stellt den Mittelwert der vier Sprachproben dar. Es wird deutlich, daß bei einer Polynomprädiktion eine Verwendung von mehr als einem Koeffizienten nicht sinnvoll ist.

Bild 8  
DPCM-Gewinn bei Polynomprädiktion.  
Sprecher 13 (TP-Sprache)

Aus Gl. 3.6.1 folgt für

N=1:  $\hat{s}(k) = p(k-1)$

Der Schätzwert ist mit dem letzten rekonstruierten Wert identisch. Die gesamte Rückkopplungsschleife entspricht einem idealen Integrierer, d.h. es ist  $\hat{s}(k) = \sum_{j=-\infty}^{k-1} g(d(k))$ .

Während bei einer optimalen Prädiktion immer ein Gewinn zu verzeichnen ist (Gl. 3.3.6 und Gl. 3.3.10),

$$h_{1*} = r_S(1) = \varrho(1) \quad ; \quad G_{S*} = (1 - \varrho^2(1))^{-1}$$

ergibt sich bei einem Polynomprädiktor nur dann ein Gewinn, wenn  $\varrho(1) > 0,5$  ist (Gl. 3.6.1 und Gl. 3.5.4 mit  $G_{Su} = G_{S*}$ ):

$$h_1 = 1 \quad ; \quad G = 2(1 - \varrho(1))^{-1}$$

N=2:  $\hat{s}(k) = 2p(k-1) - p(k-2)$

Dieser Ausdruck entspricht einer Geraden-Extrapolation. Es läßt sich leicht zeigen, daß ein System mit N=2 nur dann einem System mit N=1 überlegen ist, wenn  $3\varrho(1) > 2 + \varrho(2)$  gilt.

In beiden Beispielen wurde der Einfluß des Quantisierungsfehlers nicht berücksichtigt.

4. Spektrale Darstellung der Differenz-Pulsmodulation

Die Angabe der Spektren in der DPCM-Schleife vermittelt einige weitere Einsichten in die Funktionsweise dieser Codierung. Unter der - auch im Abschnitt 3 getroffenen - Annahme, daß der Quantisierungsfehler nicht mit dem Differenzsignal korreliert ist, ergibt sich ein Netzwerk, in das die Signale  $s(n)$  und  $q(n)$  eingespeist werden. Zunächst soll angenommen werden, daß  $s(n)$  und  $q(n)$  eine  $z$ -Transformierte besitzen. Dann gilt mit den Bezeichnungen von Bild 9 unter anderem:

$$D_Q(z) = D(z) - Q_D(z) \quad (4.1)$$

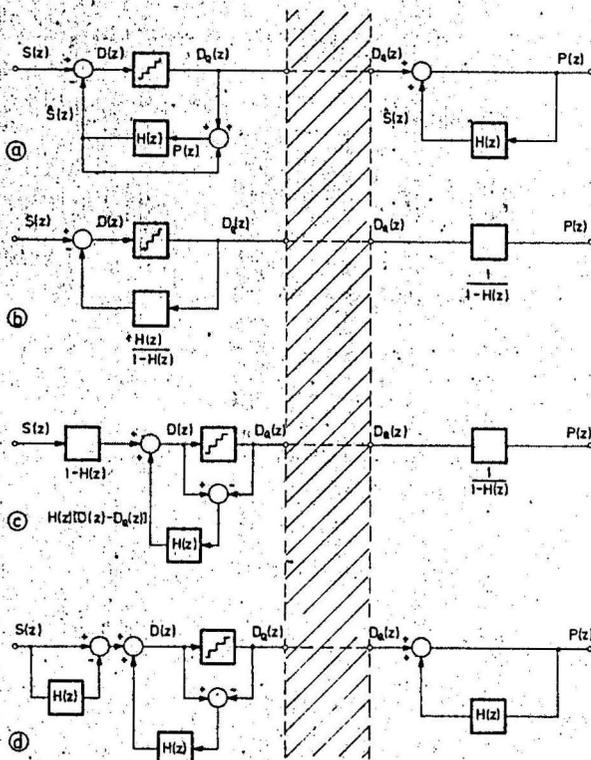


Bild 9  
DPCM-System (a) und  
äquivalente Darstellungen

Die Gleichung 4.1 könnte leicht zu dem Schluß verleiten, daß über den Kanal das Differenzsignal mit dem Spektrum  $D(z)$  und ein Quantisierungsrauschen mit dem Spektrum  $Q_D(z)$  übertragen wird.  $Q_D(z)$  ist ein weißes Spektrum, wenn die Quantisierungsfehler nicht miteinander korreliert sind; der Empfänger mit der Übertragungsfunktion  $(1-H(z))^{-1}$  (Bild 9 b) würde danach den Quantisierungsfehlern eine Korrelation aufzwingen, so daß das Quantisierungsrauschen am Ausgang der Übertragungsstrecke eine Färbung erhalten hätte. Diese Schlußfolgerung aus der Gl. 4.1 ist jedoch falsch, da auch das Spektrum  $D(z)$  einen Anteil des Quantisierungsfehlers enthält. Für das Spektrum des Differenzsignals ergibt sich

$$D(z) = S(z) - \hat{S}(z) \quad (4.2)$$

Ist  $H(z)$  die Übertragungsfunktion des Prädiktors, so folgt

$$\begin{aligned} D(z) &= S(z) - P(z) \cdot H(z) \\ &= S(z) - [D_Q(z) + \hat{S}(z)] H(z) \end{aligned}$$

Mit den Gl. 4.1 und 4.2 ergibt sich

$$\begin{aligned} D(z) &= S(z) - [D(z) - Q_D(z) + S(z) - D(z)] H(z) \\ &= S(z) [1 - H(z)] + Q_D(z) H(z) \end{aligned}$$

Damit folgt für das quantisierte Differenzsignal:

$$\begin{aligned} D_Q(z) &= D(z) - Q_D(z) \\ &= S(z) [1 - H(z)] - Q_D(z) [1 - H(z)] \quad (4.3) \end{aligned}$$

Das Ausgangssignal des DPCM-Codierers setzt sich also (additiv) aus 2 Anteilen zusammen, nämlich den jeweils mit  $(1-H(z))$  gefilterten Folgen des Eingangssignals und des Quantisierungsfehlers (Bild 9d). Sind  $\{s(n)\}$  und  $\{q(n)\}$  zeitdiskrete stochastische Prozesse mit den Leistungsdichtespektren  $\phi_S(z)$  und  $\phi_Q(z)$ , so ergibt sich für das Leistungsdichtespektrum des Ausgangs

$$\phi_{D_Q}(z) = \phi_S(z) |1 - H(z)|^2 + \phi_Q(z) |1 - H(z)|^2$$

Wie man leicht zeigen kann, sind die Kreuzleistungsdichten zwischen den beiden Anteilen Null.

Bei optimaler Auslegung des Prädiktors wird im DPCM-Codierer eine Dekorrelation des Eingangssignals vorgenommen; das über den Prädiktionszweig zurückgeführte Quantisierungsrauschen wird jedoch korreliert. Der Decodierer bewirkt umgekehrt eine Korrelation der Anteile; damit wird aber die Färbung des Quantisierungsrauschens wieder rückgängig gemacht, da der Decodierer eine Multiplikation der Eingangs-Leistungsdichtespektren mit  $|1-H(z)|^{-2}$  bewirkt. Dem Ausgangssignal ist also wie bei der PCM ein weißes Rauschen überlagert.

Bild 10 zeigt die bei Simulationsläufen von PCM- und DPCM-Systemen gemessenen Autokorrelationsfunktionen des Differenzsignals  $\{d(k)\}$ . Es wird deutlich, daß das Differenzsignal bei zunehmender Zahl der Prädiktionskoeffizienten weitgehend dekorreliert wird. Wie die Messungen zeigten, ist das Quantisierungsrauschen, d.h. die Differenz zwischen den Abtastwerten des Eingangssignals und des rekonstruierten Signals, vollkommen dekorreliert.

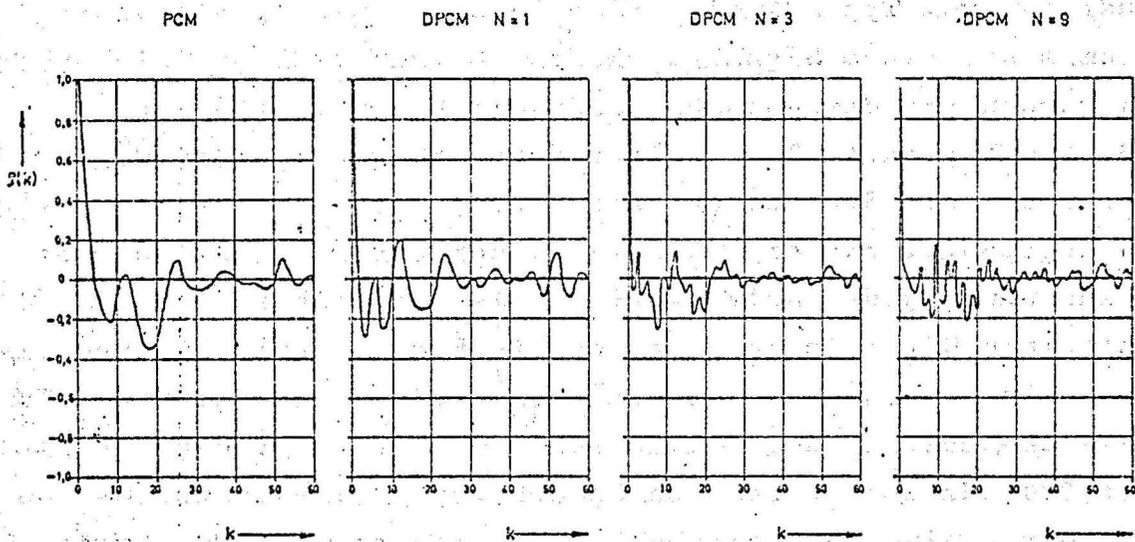


Bild 10

Normierte Autokorrelationsfunktionen (1,25 ms-Raster) des Quantisierereingangssignals verschiedener Systeme Sprecher 13 (TP-Sprache)

Ist eine Färbung des Quantisierungsrauschens erwünscht, so können die Strukturen nach Bild 9 c bzw. Bild 9 d verwendet werden. Wie aus Gl. 4.3 hervorgeht, sind diese Strukturen mit der DPCM-Struktur identisch. Eine Färbung des Quantisierungsrauschens wird erreicht, wenn das Rückführungsfiter für das Quantisierungsrauschen entsprechend geändert wird (Technischer Bericht in Vorbereitung).

## 5. Analyse der verwendeten Sprachsignale

### 5.1 Aufnahme der Sprachsignale

Als Quelle der Simulationsprogramme diente natürliche, flüssig gesprochene Sprache von 2 männlichen und 2 weiblichen Sprechern (männliche Sprecher: Nr. 10 und 13; weibliche Sprecher: Nr. 12 und 15). Die Sprachproben wurden in einem schalltoten Raum mit einem Kondensatormesßmikrophon auf Tonband aufgezeichnet; der Mikrophonabstand betrug 1 m. Als Bandgerät diente ein Revox-Tonbandgerät vom Typ A77 mit einer Bandgeschwindigkeit von 19.05 cm/s und einem Signal-Geräusch-Abstand  $\geq 58$  dB. Die Sprachproben wurden vor dem Aufnahmegerät mit einem Tiefpaß auf 3.4 kHz bandbegrenzt. Zur Weiterverarbeitung auf einem Digitalrechner wurde die Sprache auf dem Hybrid-Rechensystem CAE 90-40/RA 700 abgetastet und digitalisiert. Bei einer Abtastfrequenz von 8 kHz wurde jeder Abtastwert mit 14 bit linear quantisiert und auf einem Digitalband gespeichert. Für die Analysen und Simulationsläufe wurde uns von dem Fachbereich Kybernetik der TU Berlin Rechenzeit an der Anlage IBM 360/67 zur Verfügung gestellt. Die auf 3.4 kHz bandbegrenzten Sprachproben (TP-Sprache 0-3400 Hz) konnten durch eine zusätzliche Hochpaßfilterung auf den Bereich

von 300 Hz - 3400 Hz begrenzt werden (HP-Sprache). Als Hochpaß diente ein rekursives digitales Filter.

### 5.2 Bestimmung der Autokorrelationsfunktion

Zur Ermittlung der Autokorrelationsfunktion (AKF) der mit 8 kHz abgetasteten Sprache wurden Sprachproben von 55 sec Dauer verwendet. Die AKF wurde als Fourier-Transformierte der Abschätzung des Leistungsdichtespektrums bestimmt. Dazu wurde die Sprache in aufeinanderfolgende Segmente von jeweils  $1024=2^{10}$  Abtastwerten zerlegt. Jedes Segment wurde mit einer Hamming-Fensterfunktion /15/ bewertet und mit Hilfe des FFT-Algorithmus /16/ fouriertransformiert. Der Mittelwert der Periodogramme aller Segmente bildet die Abschätzung des Leistungsdichtespektrums /17/. Die Bilder 11 und 12 zeigen den Verlauf der normierten, d.h. auf die mittlere Leistung bezogenen Autokorrelationsfunktion für TP-Sprache (0-3400 Hz) und BP-Sprache (300 - 3400 Hz).

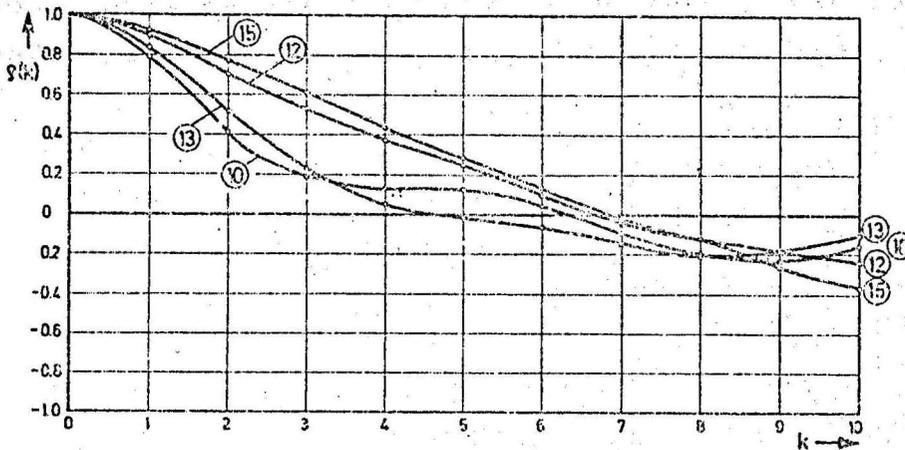


Bild 11

Normierte Autokorrelationsfunktionen (125  $\mu$ s-Raster) für 4 verschiedene Sprecher.

TP-Sprache

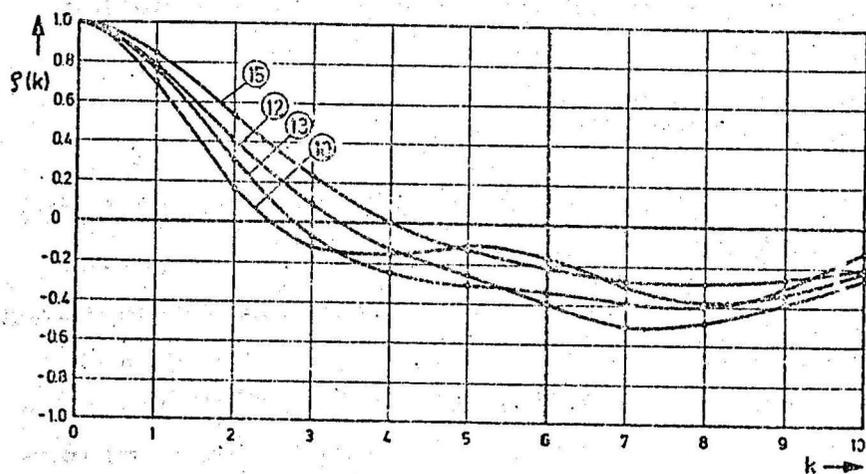


Bild 12  
Normierte Autokorrelationsfunktionen  
(125  $\mu$ s-Raster) für  
4 verschiedene Sprecher  
BP-Sprache

### 5.3 Bestimmung der Amplitudendichteverteilung

Zur Ermittlung der Amplitudendichteverteilung wurden Sprachproben von 17 sec Dauer verwendet. Es wurde die Häufigkeitsverteilung über  $s(k)$  bei einer Klasseneinteilung von  $s(k)/8$  berechnet;  $s$  ist der auf den Effektivwert  $\sqrt{S}$  bezogene Augenblickswert. Die Verteilungsdichtefunktion ergibt sich als Ableitung der Häufigkeitsverteilungsfunktion. Der in Bild 13 gezeigte Verlauf

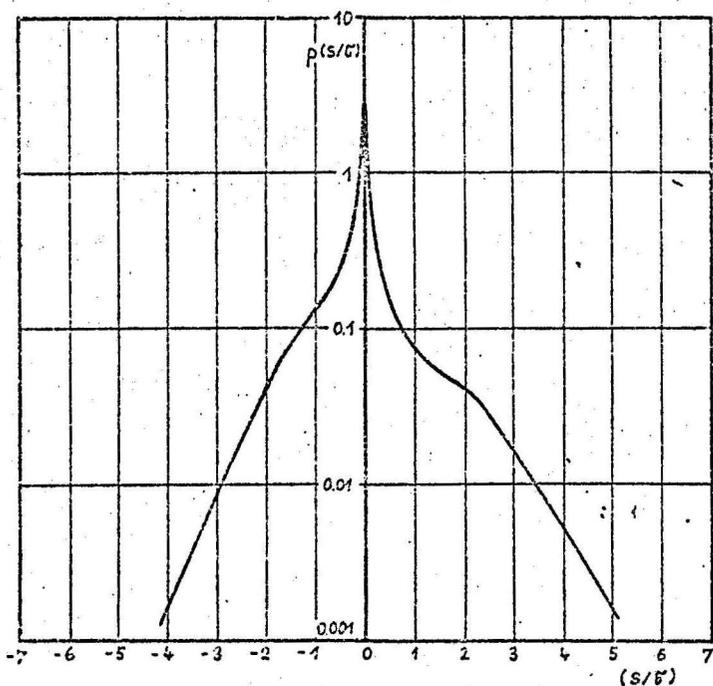


Bild 13  
Amplitudendichteverteilung, über 4 Sprecher gemittelt (TP-Sprache)

wurde aus dez. Verteilungen der TP-Sprache von 4 Sprechern gemittelt. Bei BF-Sprache ergibt sich ein verringerter Streubereich /18/. Die von Weiß /19/ erwähnte Unsymmetrie ist hier sehr stark ausgeprägt. Die Unterschiede zu den Messungen von Weiß /19/ dürften von den bedeutend kürzeren Analysedauern unserer Messungen herrühren.

## 6. DPCM-Sprachcodierung

### 6.1 Bisher bekannt gewordene Ergebnisse

Während über die Verwendung der DPCM für Bildsignale eine grosse Zahl von Untersuchungen vorliegen (s.z.B. /20,21,22/), sind über die Differenz-Pulscodemodulation von Sprachsignalen bisher nur wenige Ergebnisse bekannt geworden. Von Nitadori /23/ stammt eine theoretische Untersuchung eines DPCM-Systems mit einem Prädiktorkoeffizienten ( $N=1$ ). Bei den Berechnungen wurde eine Exponentialverteilung der Amplitudendichte des Sprachsignals vorausgesetzt. Die bisher ausführlichste Untersuchung von McDonald /13/ basiert auf Ergebnissen von Rechnersimulationen. Das verwendete DPCM-System hatte nur einen nichtoptimalen Prädiktorkoeffizienten (Polynomprädiktion:  $h_1 = 1$ ;  $h_j = 0 \forall j \geq 2$ ). Die Simulation eines 4-bit-Systems ergab einen Gewinn von etwa 7 dB; dabei wurde mit 9.6 kHz abgetastete Sprache als Eingangssignal benutzt. Nach Umrechnung auf eine Abtastfrequenz von 8 kHz ergibt sich ein Gewinn von 5.7 dB.

Ergebnisse über den subjektiv empfundenen Gewinn der DPCM-Codierung gegenüber einer Pulscodemodulation liegen zwar vor, ergeben aber kaum übereinstimmende Aussagen. Aus subjektiven Vergleichs-Untersuchungen von Donaldson et.al. /24,25/ geht hervor, daß ein 56-kbit-PCM-System einem 40-kbit-DPCM-System äquivalent ist, wenn das Sprachsignal auf 3.4 kHz begrenzt wird und eine Prädiktion mit einem vorhergehenden Abtastwert vorgenommen wird ( $N=1$ ;  $h_1 = 1$ ;  $h_j = 0 \forall j \geq 2$ ). Es entsteht somit ein Gewinn von 2 bit je Abtastwert. Neuere Ergebnisse von Chan und Donaldson /26/

erfordern eine Korrektur der in /24/ und /25/ gemachten Aussagen: aus der Abb. 11 b in /26/ läßt sich für eine Abtastfrequenz von 8 kHz bei genügend feiner Quantisierung ein Gewinn von etwa 6.3 dB ablesen. Cattermole /12/ berichtet, daß der Gewinn unter 1 bit/Abtastwert sinkt, wenn statt eines hochwertigen Mikrophons ein Kohlemikrofon verwendet wird. Auch Closs /18/ kommt zu dem Ergebnis, daß ein Gewinn von 1 bit/Abtastwert erreichbar ist, wenn man "eine sehr geringfügige Verzerrung in Kauf" nimmt.

## 6.2 Motivation der eigenen Untersuchungen

Die bisher bekannt gewordenen Ergebnisse lassen keine eindeutige Aussage über den Nutzen einer DPCM-Codierung von Sprachsignalen für eine Nachrichtenreduktion zu:

- a) Die in /13/, /24/, /25/ und /26/ mitgeteilten Ergebnisse stützen sich auf Analysen von Sprachproben eines Sprechers. Unsere Untersuchungen erfassen Sprachproben von vier Sprechern (2 männlich, 2 weiblich).
- b) Die unter Punkt a) genannten Veröffentlichungen stützen sich auf Analysen einer Sprachprobe von jeweils höchstens 5 sec. Dauer. Bei unseren Simulationsläufen wurden Sprachproben von 17 sec Dauer benutzt; zur Bestimmung der statistischen Kenndaten wurden jeweils Sprachproben von 55 sec Dauer analysiert.
- c) Bisherige Ergebnisse basieren auf Sprachsignalen, die keiner Hochpaßfilterung unterworfen waren. Unsere Ergebnisse erfassen auch Sprachsignale, die auf den Frequenzbereich von 300 Hz bis 3400 Hz bandbegrenzt sind (BP-Sprache). Bei der Verwendung der DPCM in öffentlichen Fernsprechnetzen kann nicht vorausgesetzt werden, daß der Frequenzanteil unterhalb von 300 Hz noch im zu codierenden Signal enthalten ist. Es zeigt sich aber, daß dieser Frequenzanteil für die Prädiktion von großer Bedeutung ist (s. Abschnitt 6.3.2).

- d) Statt nur den letzten rekonstruierten Wert zur Vorhersage des nächsten Eingangswertes heranzuziehen /13,26/, wurde ein Prädiktor verwendet, der eine Bestimmung des Schätzwertes aus bis zu neun Rekonstruktionswerten ermöglicht.
- e) Eine analoge Realisierung eines mehrstufigen Prädiktors ist aus Gründen der zu fordernden Genauigkeit der Vorhersage sowie aus Stabilitätsgründen oft nicht möglich; Donaldson /25/ berichtet, daß es nicht möglich war, den benutzten Codierer auf eine Vorhersage mit Ausnutzung von 2 rekonstruierten Werten zu erweitern. Es wurde daher untersucht, wie die Güte eines DPCM-Systems durch eine digitale Realisierung eines Teils des DPCM-Netzwerkes beeinflusst wird. Aus den in /7/ veröffentlichten Ergebnissen geht hervor, daß die Nachkommastellen der Prädiktionskoeffizienten mit maximal 4 bit dargestellt werden können, ohne daß der Gewinn abnimmt.
- f) Eine adaptive Anpassung des Prädiktors an die sich ändernden statistischen Kenngrößen des Eingangssignals kann den Gewinn gegenüber einer nichtadaptiven DPCM-Codierung erhöhen. Die vorliegenden Ergebnisse können als Vergleichsdaten für eine solche adaptive Codierung dienen. Eine Veröffentlichung über adaptive Differenz-Pulscodemodulation von Sprachsignalen ist in Vorbereitung.

### 6.3 Simulationsergebnisse

#### 6.3.1 Signal-Rausch-Abstände

Die Simulation wurde mit Sprachproben von 4 verschiedenen Sprechern (2 weiblich, Nr. 12,15 und 2 männlich, Nr. 10,13) durchgeführt. (Abtastfrequenz 8 kHz; Bandbegrenzung wahlweise 0-3400 Hz (TP-Sprache) oder 300 - 3400 Hz (BP-Sprache); Dauer der Sprachproben: jeweils 17 sec). Die Prädiktorkoeffizienten der DPCM-Systeme wurden gemäß Gl. 3.3.7 für jeden Sprecher aus der AKF berechnet, die über 55 sec Dauer ermittelt wurde (vgl. Abschnitte 3.3 bzw. 5.2).

Bei der digitalen Übertragung von Sprachsignalen über PCM- oder DPCM-Systeme treten zwei Arten von nichtlinearen Verzerrungen auf:

- a) ein über der Frequenz gleichverteiltes granulares Quantisierungsrauschen
- b) ein durch die Übersteuerung des Quantisierers hervorgerufenes Begrenzungsrauschen

Zu a) Quantisierungsrauschen entsteht, wenn die Eingangsamplitude kleiner als der Aussteuerungsbereich des Quantisierers ist ( $s(k) \leq s_c$  nach Bild 3); es hängt ab von der Stufenhöhe der einzelnen Quantisierungsstufen, bei gegebenem Aussteuerbereich also von der Zahl der Quantisierungsstufen (und der Art der Quantisierercharakteristik  $g(\cdot)$ ). Eine Verdopplung der Zahl der Quantisierungsstufen bei gleichem Aussteuerungsbereich (und gleichem Eingangssignal) verbessert den Signal-Rausch-Abstand um 6 dB. Grundsätzlich ist es möglich, bei vorgegebener Amplitudenverteilung des Eingangssignals eine Quantisierercharakteristik  $g(\cdot)$  anzugeben, die zu einem maximalen Signal-Rausch-Abstand führt /9,10/. Bei Sprachübertragung ist jedoch eine logarithmische bzw. quasilogarithmische Quantisierercharakteristik notwendig, um sicherzustellen, daß auch bei geringer Aussteuerung des Quantisierers das Signal-Rausch-Verhältnis einen Grenzwert nicht unterschreitet (/11,12/, s. auch Abschnitt 3.2). Bei den Simulationen wurden logarithmische Kennlinien mit einem Kennwert  $\mu = 100$  benutzt /11/.

Zu b) Begrenzungsrauschen entsteht bei der Übersteuerung des Quantisierers ( $s(k) > s_c$  nach Bild 3). Seine Größe hängt vom Aussteuerungsbereich des Quantisierers und der Verteilungsdichtefunktion des Eingangssignals ab. Im Unterschied zum granularen Quantisierungsrauschen wird das Begrenzungsrauschen nicht so störend empfunden.

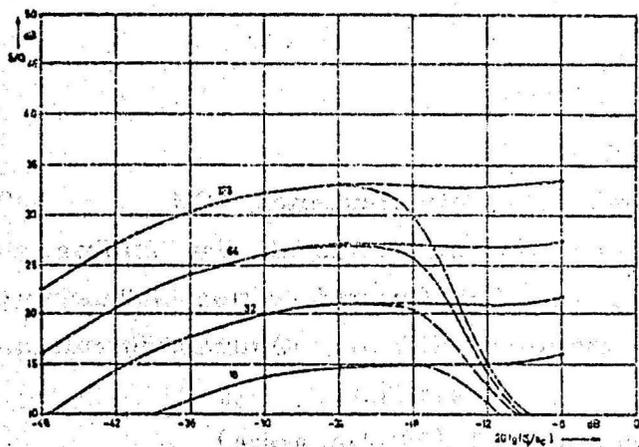


Bild 14  
Simulationsergebnisse: PCM;  $\mu = 100$ .  
Signal-Rausch-Abstände in Abhängigkeit von der Quantisierer-Belastung,  
Parameter: Zahl der Quantisierungsstufen.

Sprecher 13 (TP-Sprache)

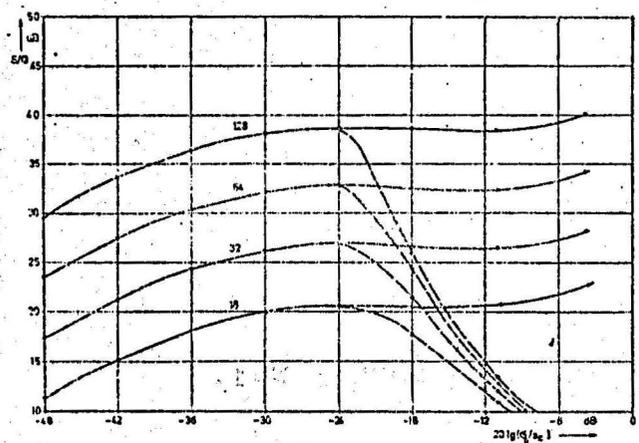


Bild 15  
wie 14, aber DPCM ; N=1.

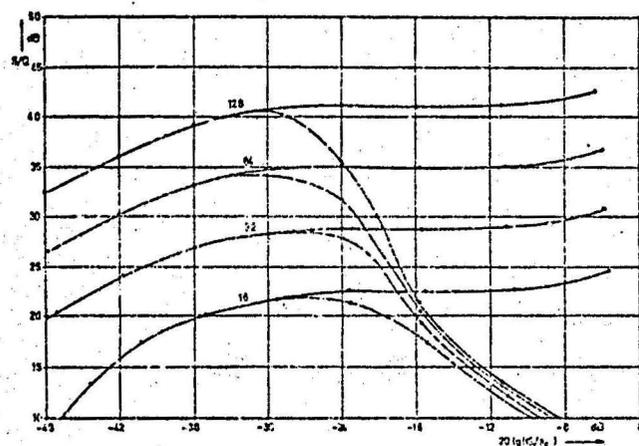


Bild 16  
wie 14, aber DPCM ; N=3.

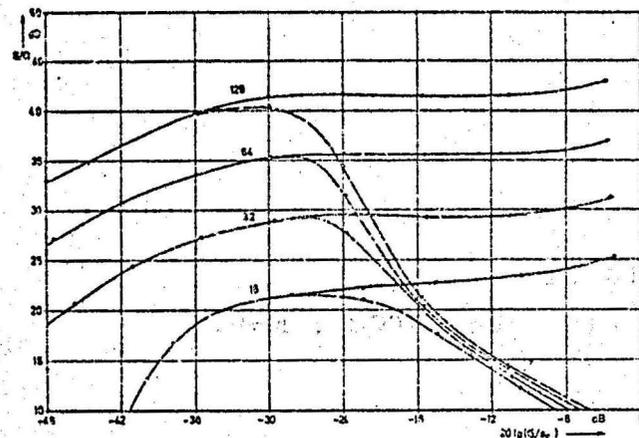


Bild 17  
wie 14, aber DPCM ; N=9.

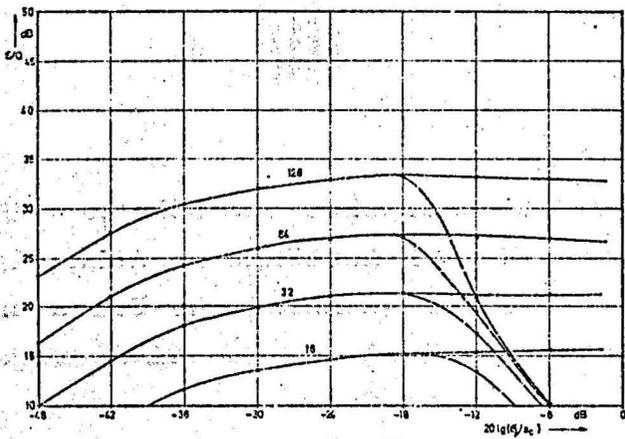


Bild 18  
Simulationsergebnisse: PCM;  $\mu = 100$ .  
Signal-Rausch-Abstände in Abhängig-  
keit von der Quantisierer-Belastung.  
Parameter: Zahl der Quantisierungs-  
stufen.

Sprecher 12 (TP-Sprache)

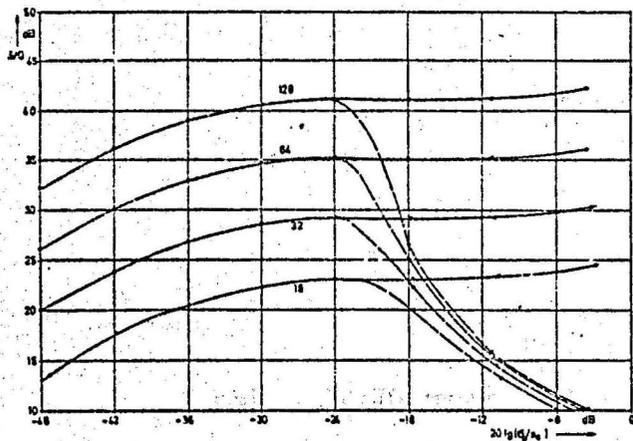


Bild 19  
wie 18, aber DPCM ;  $N=1$ .

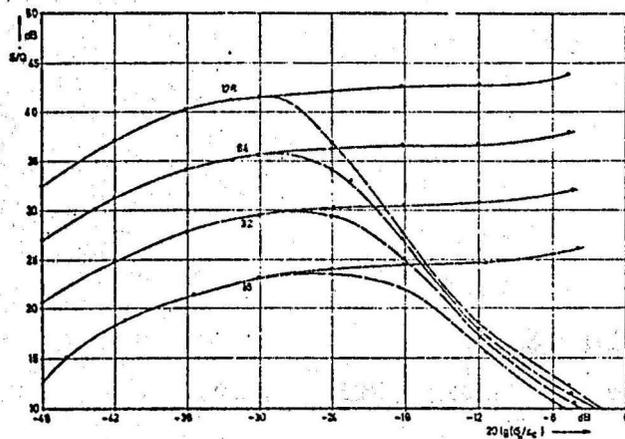


Bild 20  
wie 18, aber DPCM ;  $N=3$ .

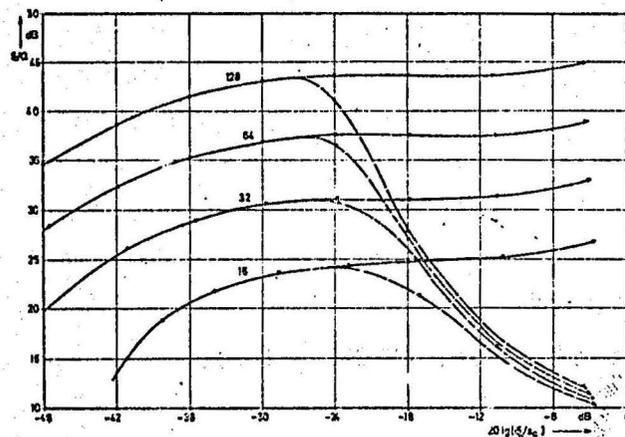


Bild 21  
wie 18, aber DPCM ;  $N=9$ .

Die Bilder 14 bis 21 zeigen den Signal-Rausch-Abstand eines PCM-Systems und mehrerer DPCM-Systeme mit verschiedenen Prädiktoren als Funktion des Aussteuerungsbereichs des Quantisierers; der Parameter gibt die Zahl der Quantisierungsstufen an. Die Kurven 14 - 17 bzw. 18 - 21 gelten für Sprachproben eines männlichen bzw. eines weiblichen Sprechers; der Frequenzbereich beträgt 0-3400 Hz. Die gestrichelten Verläufe geben den Einfluß der Begrenzung wieder.

Bei der Pulsmodulation wird das Begrenzungsrauschen erst bei Pegeln oberhalb von etwa  $20 \lg (G/s) = -(14 \cdot 0.16)$  dB wahrnehmbar, wobei eine Grenze von  $-(7 \cdot 0.9)$  dB durchaus akzeptiert werden kann /27/. Bei DPCM-Systemen müßte das Begrenzungsrauschen bei gleichen Belastungen wie bei dem PCM-System sichtbar werden, wenn die Amplitudendichteverteilungen des Eingangs- und Differenzsignals gleich sind. Die Analyse der Übersteuerungswahrscheinlichkeit des Quantisierers weist - zumindest für grobe Augenblickswerte -, auf eine Änderung der Dichteverteilungen hin. Bild 22 zeigt die Wahrscheinlichkeit für das Auftreten eines Quantisierer-Eingangssignals, das den Aussteuerungsbereich

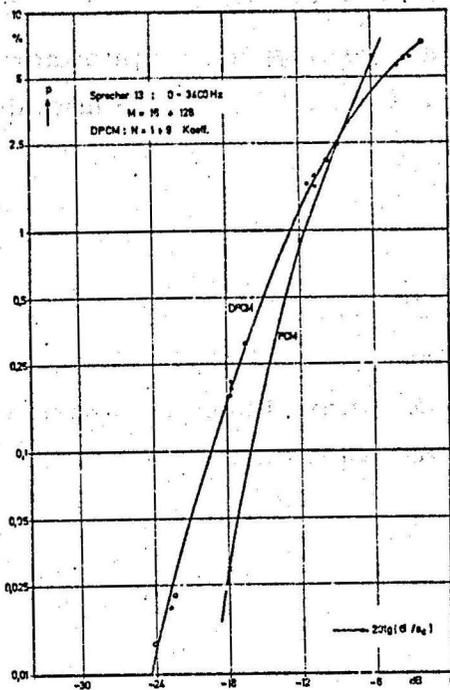


Bild 22  
Quantisierer-Übersteuerungswahrscheinlichkeit

des Quantisierers überschreitet ( $s > s_c$ ). Es zeigt sich, daß bei Belastungen unter ca. - 10 dB die Übersteuerungswahrscheinlichkeit bei DPCM in jedem Fall höher ist als bei PCM. Das Begrenzungsrauschen wird somit bei DPCM schon bei kleineren Pegeln sichtbar (etwa - 24 %/ 30). dB bei DPCM; PCM: - 20 dB, vgl. Bilder 14 - 21). Für die Wahrnehmbarkeit dieses Rauschens bei DPCM-Systemen liegen keine genaueren Angaben vor. Die Ursachen (und auch die subjektiven Eindrücke) des Begrenzungsrauschens sind bei PCM und DPCM verschieden. Bei der Pulscode-modulation führt eine Begrenzung auf eine Abkappung von Signalanteilen des Sprachsignals mit großer Amplitude und damit wegen des quasi-periodischen Auftretens dieser Amplitudenmaxima zu harmonischen Verzerrungen. Bei der DPCM treten große Differenzsignale bei einer schlechten Vorhersage auf, also z.B. bei starken Steigungen des Spracheingangssignals (Explosivlaute etc). Der eigene subjektive Eindruck ist der, daß noch höhere Fehlerleistungen als bei der PCM-Amplitudenbegrenzung zugelassen werden können, da die Sprache durch Übersteuerung zwar dumpfer klingt, ohne daß jedoch der Eindruck einer Verzerrung entsteht.

Die Bilder 23 und 24 zeigen den Signal-Rausch-Abstand als Funktion der Quantisiererbelastung für das PCM-System und die 3 untersuchten DPCM-Systeme. Die Ergebnisse gelten für Sprecher 13, die Quantisiererwortlänge beträgt 7 bzw. 4 bit. Mit zunehmender

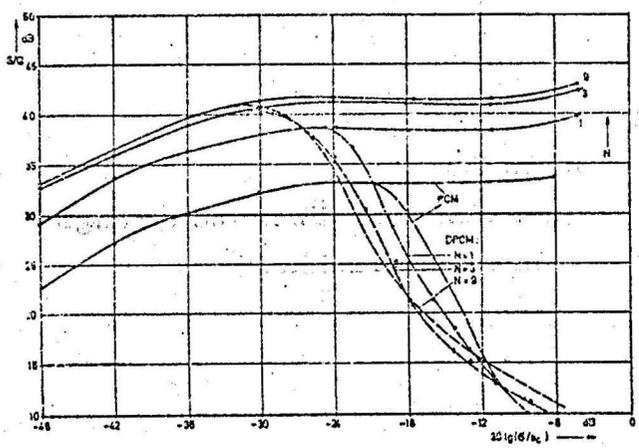


Bild 23  
Vergleich PCM-DPCM  
128 Quantisierungsstufen  
Sprecher 13 (TP-Sprache)

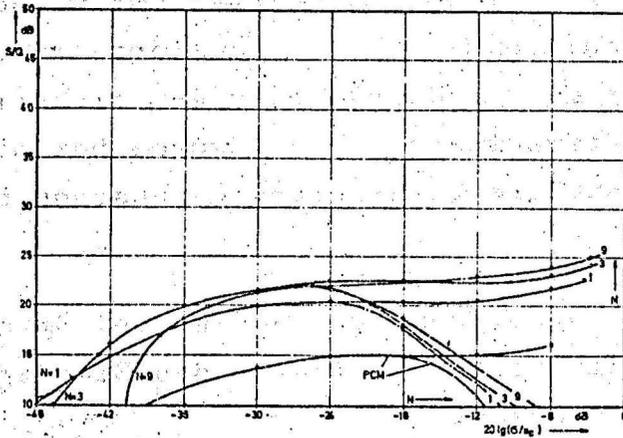


Bild 24  
Vergleich PCM-DPCM  
32 Quantisierungsstufen  
Sprecher 13 (TP-Sprache)

Zahl der Prädiktorkoeffizienten verbessert sich der Signal-Rausch-Abstand bei großen Wortlängen stetig. Bei kleinen Wortlängen nimmt der Signal-Rausch-Abstand mit abnehmender Belastung und steigender Zahl der Prädiktorkoeffizienten immer rascher ab (Bild 24). Von den möglichen 16 Quantisiererstufen können durch extrem niedrige Aussteuerung nur einige wenige Stufen auftreten, so daß große Quantisierungsfehler auftreten. Das DPCM-System enthält eine Rückkopplungsschleife, die naturgemäß auch instabil werden kann. Wird ein sehr großer Quantisierungsfehler zurückgeführt, wird die Wahrscheinlichkeit für eine schlechte Vorhersage größer, je mehr (stark fehlerbehaftete) vergangene Abtastwerte im Prädiktor gespeichert werden. Die Gewinne nehmen daher im Bereich extrem niedriger Belastung mit zunehmender Zahl der Prädiktorkoeffizienten ab.

### 6.3.2 DPCM-Gewinne

Der Gewinn eines DPCM-Verfahrens gegenüber einer PCM-Codierung ist das Verhältnis der Varianzen der Quantisierungsfehler:

$$G_D = \frac{\sigma_P}{\sigma_D} \quad (2.3.1)$$

Dieser Gewinn konnte bei unseren Simulationen bestimmt werden, da neben den DPCM-Messungen auch PCM-Systeme simuliert wurden. Da der Einfluß der Begrenzung in der Wahrnehmbarkeit sicherlich bei Aussteuerungen  $s_c/6 \geq 4$  bedeutungslos ist, wurde bei der Bestimmung von  $G_D$  nur das granulare Quantisierungsrauschen berücksichtigt.

Wie im Abschnitt 3.2 gezeigt wurde, kann der Gewinn bei Sprachsignalen auch durch das Verhältnis der Varianzen des Eingang- und Differenzsignals bestimmt werden.

$$G_S = \frac{S}{S_D} \quad (3.2.1)$$

$G_D$  bzw.  $G_S$  können nun mit den theoretischen, aus den Langzeit-AKF-Werten berechneten Gewinnen verglichen werden:

$$G_{Su} = \frac{1 - e_D^2 \underline{h}_*^T \underline{R}_Q \underline{h}_*}{1 - \underline{h}_*^T \underline{R}_S \underline{h}_*}$$

$G_{Su}$  ist der suboptimale Gewinn, das ist der maximal erreichbare Gewinn, wenn der  $\underline{h}$ -Vektor der Einstell-Koeffizienten aus den statistischen Kenngrößen des Eingangsprozesses bestimmt wird ( $\underline{h} = \underline{h}_*$ , s.Gl. 3.3.6). Wird der Einfluß der Fehlerrückkopplung nicht berücksichtigt ( $e_D^2 \rightarrow 0$ ), so ergibt sich der maximal mögliche Gewinn

$$G_{S*} = G'_{opt} = \frac{1}{1 - \underline{h}_*^T \underline{R}_S \underline{h}_*} \quad (3.3.10)$$

Die Bilder 25 und 26 zeigen den theoretischen Gewinn  $G'_{opt} = G_{S*}$ , der über vier Sprecher gemittelt wurde, in Abhängigkeit von der Zahl der Prädiktorkoeffizienten. Bild 25 gilt für TP-Sprache, Bild 26 für BP-Sprache; der schraffierte Bereich gibt den Streubereich an.

... (mirrored text from the reverse side of the page) ...

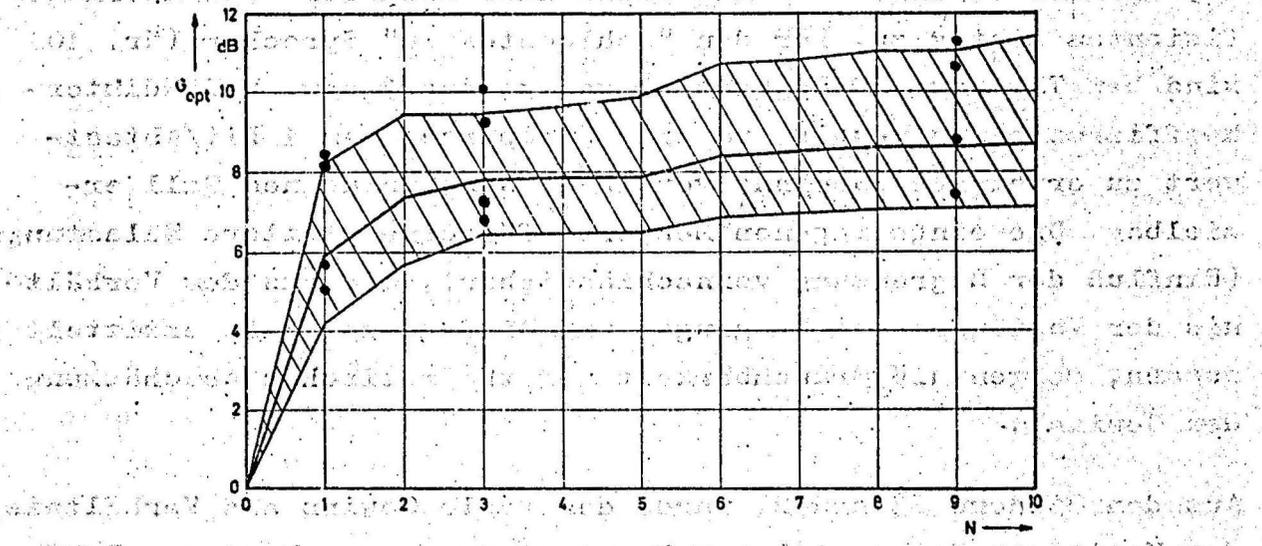


Bild 25

DPCM-Gewinn nach Gl. (3.3.10)

TP-Sprache

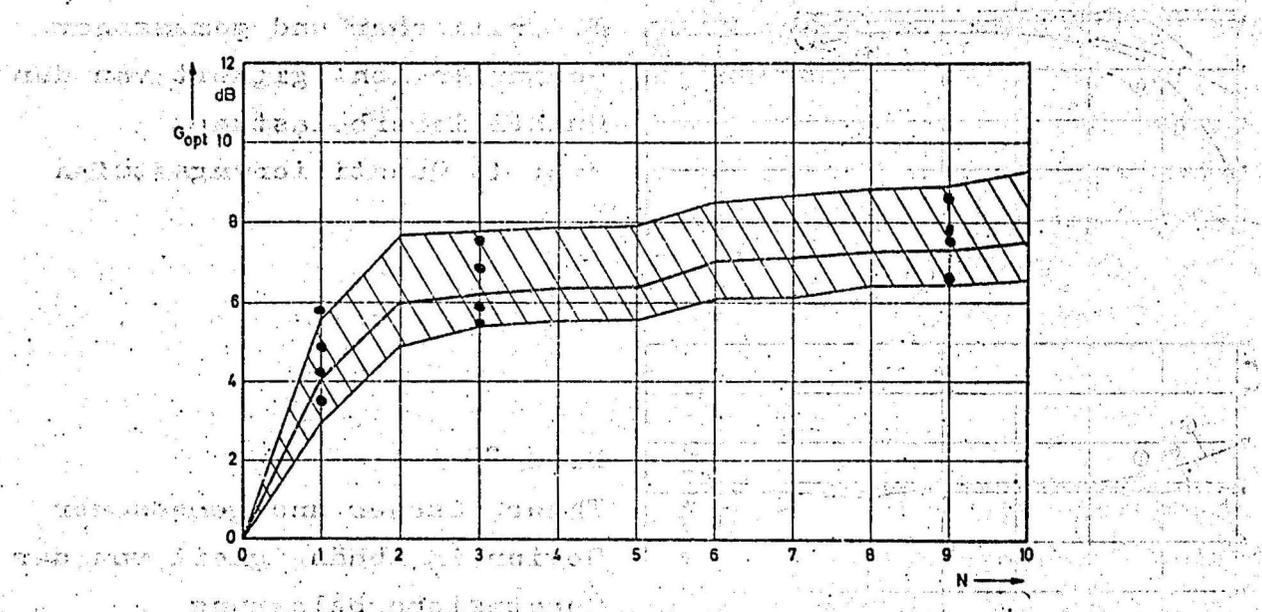


Bild 26

DPCM-Gewinn nach Gl. (3.3.10)

BP-Sprache

Die Ausfilterung des tieffrequenten Sprachanteils zwischen 0 und 300 Hz (Wegfall der Sprachgrundfrequenz) mit recht hohem Energieanteil bewirkt eine mittlere Reduktion des Gewinns um 1.5 dB. Der Gewinn nimmt mit wachsender Zahl der Prädiktorkoeffizienten stetig zu. Für den "schlechtesten" Sprecher (Nr. 10) sind bei TP- bzw. BP-Begrenzung wenigstens 3 bzw. 6 Prädiktorkoeffizienten notwendig, um eine Einsparung von 1 bit/Abtastwert zu erreichen. Gewinne von 12 dB sind in keinem Fall erzielbar. Die eingetragenen Meßwerte für eine mittlere Belastung (Einfluß der Begrenzung vernachlässigbar), die aus dem Verhältnis der Varianzen des Eingangs- und Differenzsignals ermittelt wurden, zeigen die Brauchbarkeit der theoretischen Abschätzung des Gewinns.

Aus den Bildern 23 und 24 wurde der reale Gewinn als Verhältnis der Varianzen des granularen Quantisierungsrauschens von DPCM zu PCM ermittelt (Gl. 2.3.1; Kurve (3) in Bild 27 und 28). Die

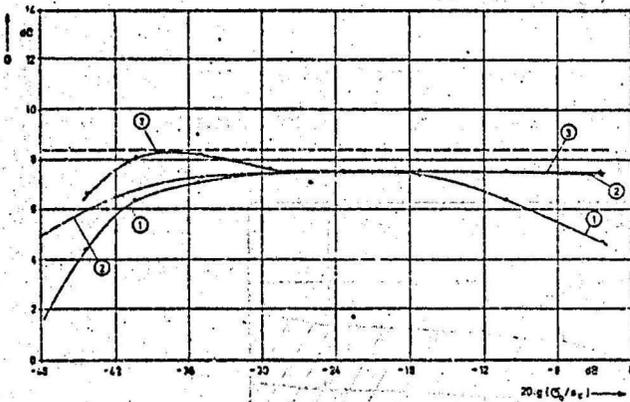


Bild 27  
Theoretischer und gemessener Gewinn in Abhängigkeit von der Quantisiererbelastung  
 $N=3$ ; 16 Quantisierungsstufen

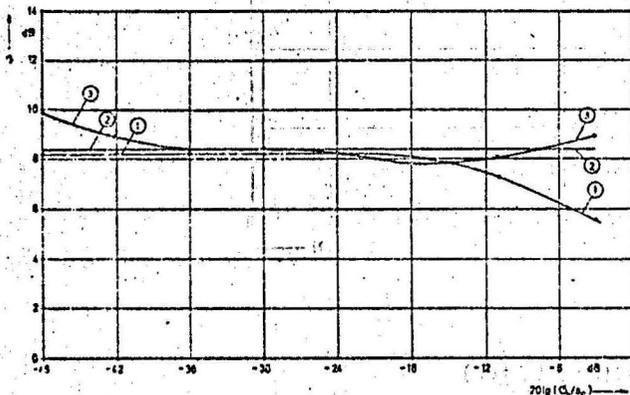


Bild 28  
Theoretischer und gemessener Gewinn in Abhängigkeit von der Quantisiererbelastung  
 $N=3$ ; 128 Quantisierungsstufen

Kurven (1) und (2) geben den als Verhältnis der Varianzen des Eingang- und Differenzsignals ermittelten Gewinn (nach Gl. 3.2.1) und den theoretischen Gewinn  $G_{Su}$  gemäß Gleichung 3.3.8 wieder. Die gestrichelte Gerade im Bild 27 stellt den Gewinn  $G'_{opt}$  nach Gl. 3.3.10 dar. Die Kurven gelten für TP-Sprache von Sprecher 13, die Zahl der Prädiktorkoeffizienten beträgt 3. Die Unterschiede sind in den verschiedenen Amplitudendichteverteilungen des Eingang- und Differenzsignals und in nicht berücksichtigten nichtlinearen Abhängigkeiten zwischen Abtastwerten begründet.

### 6.3.3 Weitere Untersuchungen

In weiteren Simulationsläufen wurde der Einfluß der Quantisiererkennlinie auf die Signal-Rausch-Abstände untersucht. Die logarithmische  $\mu$ -Kennlinie (s. Abschnitt 7) wurde beibehalten, der Grad der Kompression  $\mu$  wurde geändert. Erhebliche Abweichungen von den gezeigten Ergebnissen treten nur bei extrem kleiner Aussteuerung des Quantisierers auf, bei einer mittleren Aussteuerung ( $s_c = 16 \cdot 6$ ) verbessern sich die Signal-Rausch-Abstände bei linearer Quantisiererkennlinie ( $\mu = 0$ ) um weniger als 1 dB gegenüber logarithmischer Kompression mit  $\mu = 100$ .

Zur Verringerung des Ruhegeräusches wurde der aus Stabilitätsgründen verwendete midriser-Quantisierer (s. Abschnitt 7) gegen einen midtrear-Quantisierer ausgetauscht. Bei der digitalen Realisierung des DPCM-Systems wurden keine Stabilitätsschwierigkeiten durch den Austausch festgestellt. Durch die logarithmische Kompression mit  $\mu = 100$  ist die kleinste Quantisiererstufe bei einer 4-bit-Übertragung schon so klein, daß der Einfluß auf den Signal-Rausch-Abstand weitgehend vernachlässigbar ist.

## 7. Beschreibung des Simulations-Systems

Das Prinzip der DPCM wurde im Abschnitt 2 bereits eingehend erläutert. Die zur Beschreibung des Simulationsprogramms notwendigen Grundlagen sollen kurz zusammengefaßt werden. Bei den Si-

mulationen wurde eine fehlerfreie Kanalübertragung vorausgesetzt. Das decodierte Ausgangssignal steht auch im Sender als Eingangsfolge  $\{p(k)\}$  des Prädiktors zur Verfügung (s. Abschnitt 2.2). Es braucht also nur die Sendeeinrichtung simuliert werden.

Für die Quantisiererkennlinie ist ein logarithmischer Verlauf vorgeschrieben (s. Abschnitt 3.2). Der Quantisierer kann aufgeteilt werden in einen Kompressor mit logarithmischer Kennlinie und einen linearen Quantisierer.

Der Prädiktor ist ein nichtrekursives Filter (Abschnitt 3.1), dessen Aufbau in Bild 4 gezeigt ist.

Der Quantisierungsfehler, definiert als Differenz des Eingangs- und Ausgangssignals des nichtlinearen Quantisierers, ist identisch mit der Differenz des System-Eingangssignals und des decodierten Ausgangssignals (s. Abschnitt 2.1).

Die Signal-Rausch-Abstände werden aus dem ungefilterten Quantisierungsrauschen bestimmt, d.h. es werden die Anteile im Frequenzbereich von 3.4 bis 4 kHz mitberücksichtigt. Ist das Quantisierungsrauschen ein weißes Rauschen (was für die untersuchten Systeme erfüllt ist), so lassen sich ohne weiteres die auf den Bereich bis 3.4 kHz bezogenen Signal-Rausch-Abstände aus den gezeigten Simulationsergebnissen bestimmen.

Im folgenden sollen Einzelheiten des Systems näher erläutert werden:

#### Der Kompressor

Die Aufteilung des nichtlinearen Quantisierers in Kompressor und linearen Quantisierer ermöglicht eine einfache Programmierung und erlaubt einen leichten Austausch der Kennlinien. Als Kompondierungskennlinie wurde die  $\mu$ -Kennlinie nach Smith /11/ gewählt:

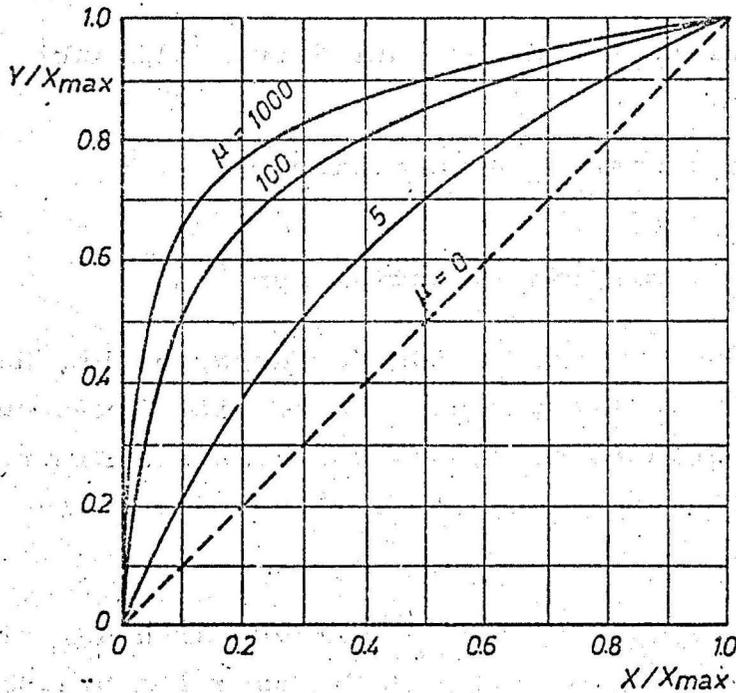


Bild 29  
logarithmische Kom-  
pressorkennlinien

Die Kennlinienschar wird durch folgende Formel beschrieben:

$$y = \frac{x_{\max} \cdot \lg \left( 1 + \mu \cdot \frac{x}{x_{\max}} \right)}{\lg (1 + \mu)} \operatorname{sign} (x)$$

Als Kompressionsgrad wurde für die Simulation der Wert  $\mu = 100$  gewählt (vgl. jedoch auch Abschnitt 6.3.3) .

#### Der lineare Quantisierer

Eine lineare Quantisierererkennlinie wurde bereits in Bild 3 gezeigt. Die Kennlinie muß jedoch nicht notwendig durch den Koordinatenursprung gehen. Eine allgemeine Kennlinie zeigt Bild 30.

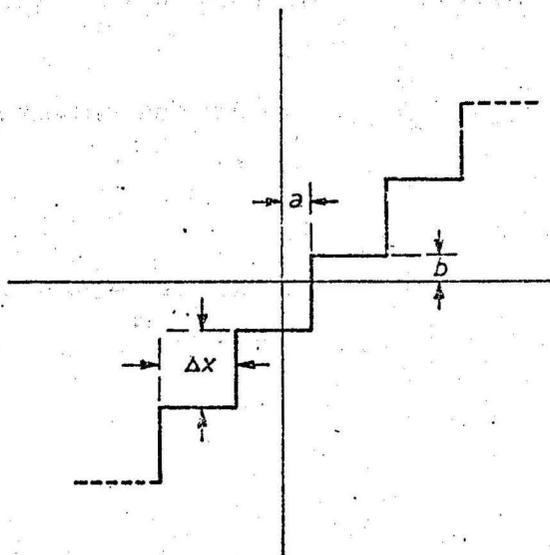


Bild 30  
Lineare Amplitudenquanti-  
sierung

Die Simulationsläufe wurden beschränkt auf 2 zum Nullpunkt symmetrische Kennlinien:

$$a = 0 ; \quad b = \frac{\Delta x}{2} \quad : \text{midriser-Quantisierer}$$

und

$$a = \frac{-\Delta x}{2} ; \quad b = 0 \quad : \text{midtreat-Quantisierer}$$

Da nur eine endliche Zahl von Stufen zur Verfügung steht, hat aus Symmetriegründen der midtreat-Quantisierer eine Stufe weniger als der midriser-Quantisierer, das Ruhegeräusch nimmt jedoch ab, da bei Eingangsspannungen um Null nur die Stufe "Null" ausgegeben wird.

Die in Abschnitt 6 gezeigten Ergebnisse beziehen sich auf ein System mit midriser-Quantisierer (s. jedoch Abschnitt 6.3.3), da nach /13/ damit ein höherer Stabilitätsgrad erreicht werden kann.

### Der DA-Wandler

Bei der DA-Wandlung der quantisierten und codierten Werte des Differenzsignals muß die logarithmische Kompondierung rückgängig gemacht werden. Bei einer Übertragung mit  $m$  bit sind  $2^m$  Amplitudenstufen beim midriser-Quantisierer bzw.  $2^m - 1$  Stufen beim midtreat-Quantisierer möglich. Aus Symmetriegründen braucht die Kennlinie nur im I. Quadranten nachgebildet werden. Alle möglichen positiven Analogwerte werden so in ein eindimensionales Feld gespeichert, daß die zu decodierende Zahl den Index des zugehörigen Matrix-Elementes angibt. Somit gilt:

$$y(i) = \frac{x_{\max}}{\mu} \cdot \left[ 10^{\frac{2i}{2^m - 1} \lg(1 + \mu)} - 1 \right] \quad \text{midtreat-Quant.}$$

$i = 1 \dots 2^m - 1$

bzw.

$$y(i) = \frac{x_{\max}}{\mu} \cdot \left[ 10^{\frac{2i}{2^m} \lg(1 + \mu)} - 1 \right] \quad \text{midriser-Quant.}$$

$i = 1 \dots 2^m$

Der Prädiktor

Der Entwurf des Prädiktors wurde bereits eingehend in Abschnitt 3 beschrieben. Die Programmierung folgt ohne weiteres aus dem in Bild 3 gezeigten Aufbau.

Dank

Die Untersuchungen wurden von der Deutschen Forschungsgemeinschaft unterstützt. Der Informatik-Forschungsbereich der Technischen Universität Berlin stellte die notwendige Rechenzeit am System IBM 360/67 zur Verfügung. Beiden Institutionen gilt unser Dank.

8. Literaturverzeichnis

- /1/ Gallager, R.G.  
Information Theory and Reliable Communication.  
J. Wiley, 1968.
- /2/ Musmann, H.G.  
Informationstheoretische Betrachtung der Nachrichtenreduktion.  
DGON-Bericht Nr. 102, Sept. 1969.
- /3/ Rupp, H.  
Codierung in der PCM-Technik.  
NTG-Fachtagung "Codierung" 1970.  
NTF 40, S.72-92.
- /4/ Richards, D.L.  
Statistical properties of speech signals.  
Proc. IEE, Vol.111 (1964), S. 941-947.
- /5/ Schlink, W.  
A redundancy reducing PCM system for speech signals.  
International Zürich Seminar on Integrated Systems for Speech, Video and Data Communications.  
Zürich, 15.-17. März 1972.
- /6/ O'Neal, J.B.  
Entropy coding in speech and television differential PCM systems  
IEEE Trans. on Inform.Theory, Vol.IT-17, S.758-761
- /7/ Noll, P.  
Sprachcodierung mit Differenz-Pulscodemodulation  
NTG-Fachbericht "PCM-Technik" (Veröffentlichung in Vorbereitung)  
Vorgetragen auf der NTG-Fachtagung München 11.-14.10., 1971.
- /8/ Noll, P.  
Nonadaptive and adaptive differential pulse code modulation of speech signals.  
Vortrag IEEE EUROCON 1971, Lausanne, 18.-22.10., 1971.
- /9/ Panter, P.F.; Dite, W.  
Quantization distortion in pulsecount modulation with nonuniform spacing of levels.  
Proc. IRE, Vol.39 (1951), S.44-48.

- /10/ Max, J.  
Quantizing for minimum distortion.  
IRE Trans. on Inform. Theory, Vol. IT-6 (1960),  
S. 7-12.
- /11/ Smith, B.  
Instantaneous companding of quantized signals.  
Bell System Techn. Journal, Vol.36 (1957),  
S. 653-709.
- /12/ Cattermole, K.W.  
Principles of pulse code modulation.  
London Iliffe Books Ltd., 1969.
- /13/ McDonald, R.A.  
Signal-to-noise and idle channel performance of  
differential pulse code modulation systems-parti-  
cular applications to voice signals.  
Bell System Technical Journal, Vol. 45(1966),  
S. 1123-1151.
- /14/ O'Neal, J.B.  
Signal-to-quantizing noise ratios for differential  
PCM.  
IEEE Trans. on Commun. Technol., Vol. COM-19 (1971),  
S. 568-570.
- /15/ Blackman, R.B.; Tukey, J.W.  
The measurement of power spectra.  
New York Dover 1959.
- /16/ Cooley, J.W.; Tukey, J.W.  
An algorithm for the machine calculation of  
complex Fourier series.  
Mathematics of Computation, Vol.19 (1965), No.297.
- /17/ Welch, P.D.  
The use of fast Fourier transform for the estimation  
of power spectra: a method based on time averaging  
over short, modified periodograms.  
IEEE Transactions on Audio and Electroacoustics,  
Vol. AU-15, 1967, S.70-73.
- /18/ Closs, F.  
Untersuchungen über das Quantisierungsrauschen bei  
der Pulscode-Modulation und dessen Verminderung.  
Dissertation Universität Stuttgart 1970, S.121.

- /19/      Weiß, W.  
Statistische Untersuchungen an Sprachschwingungen  
Nachrichtentechn. Zeitschrift 1964, S.339-348
- /20/      O'Neal, J.B.  
Predictive quantizing Systems (differential pulse  
code modulation) for the transmission of television  
signals.  
Bell System Technical Journal Vol.45 (1966),  
S. 689-721.
- /21/      Millard, J.B.; Maunsell, H.I.  
The Picturephone system: digital encoding of the  
video signal.  
Bell System Technical Journal Vol.50 (1971),  
S. 459-479.
- /22/      Habibi, A.  
Comparison of n-th order DPCM encoder with linear  
transformations and block quantization techniques.  
IEEE Trans. on Commun. Technol., Vol.COM-19,  
(1971), S. 948-956.
- /23/      Nitadori, K.  
Statistical analysis of  $\Delta$ -PCM.  
Electronics and Communications in Japan, Vol.48  
(1965), No.2 .
- /24/      Donaldson, R.W.; Chan, D.  
Analysis and subjective evaluation of differential  
pulse-code modulation voice communication systems.  
IEEE Trans. on Commun. Technology, Vol. COM-17,  
1969, S. 10-19.
- /25/      Donaldson, R.W.; Douville, R.J.  
Analysis, subjective evaluation, optimization,  
and comparison of the performance capabilities of  
PCM, DPCM,  $\Delta$ M, AM, and PM voice communication systems.  
IEEE Trans. on Commun. Technol., Vol.COM-17, 1969,  
S. 421-431.
- /26/      Chan, D.; Donaldson, R.W.  
Subjective evaluation of pre- and postfiltering in  
PAM, PCM, and DPCM voice communication systems.  
IEEE Trans. on Commun. Technol., Vol. COM-19,  
1971, S. 601-612.
- /27/      Richards, D.L.  
Transmission performance of telephone networks con-  
taining PCM links.  
Proc. IEE (London), 115(1968), S. 1245-1258

