

Mit freundlichen Grüßen:
Vom Verfasser

Springer

EIN VERFAHREN ZUR SPRACHBANDKOMPRESSION

VON DIPL.-ING. ANTON M. SPRINGER

SONDERDRUCK AUS ZEITSCHRIFT
INTERNATIONALE ELEKTRONISCHE RUNDSCHAU
17 (1963) HEFT 9, SEITEN 471—475

VERLAG FÜR RADIO-FOTO-KINOTECHNIK GMBH
BERLIN-BORSIGWALDE

Eltro GMBH & CO.
GESELLSCHAFT FÜR STRAHLUNGSTECHNIK

Ein Verfahren zur Sprachbandkompression¹⁾

Die Ausweitung des Fernsprecherverkehrs und der Nachrichtenübermittlung über große Entfernungen stellt erhebliche Anforderungen an die Technik und verlangt möglichst geringe Bandbreiten. Das beschriebene Verfahren zeigt einen Weg zur Sprachbandkompression auf völlig neuer Basis und unterstreicht daraufhin die Möglichkeit, kostspielige Filter durch Nachrichtenspeicher zu ersetzen. Zum Schluß wird ein Weg zur besseren Geheimhaltung der Nachricht auf dem Übertragungsweg aus diesem Verfahren hergeleitet.

Das Problem, die menschliche Sprache zu übertragen, wird als gelöst betrachtet. Das Problem jedoch, mit möglichst wenig Bandbreite möglichst viel Information zu übertragen, ist zwar wissenschaftlich weit erforscht, doch findet die Technik immer wieder neue Wege. Einer von diesen soll hier beschrieben werden.

Die Sprachübertragung ist nicht nur ein technisches, sie ist auch ein physiologisches Problem. Die Eigenschaften des menschlichen Ohres können in vieler Hinsicht noch nicht ganz erfaßt werden. Gelingt es aber, die Kombinationsmöglichkeiten des Ohres vollständig auszunutzen, dann ergeben sich Möglichkeiten, die Frequenzbänder des Signals nach neuen Gesichtspunkten zu beschneiden. So ist beispielsweise bekannt, daß unser Gehör Phasensprünge — soweit sie nicht zu häufig in einer Zeiteinheit auftreten — nicht wahrnimmt und daß tiefe Frequenzen — etwa die des Grundtons der Sprache — zur Silbenverständlichkeit wesentlich weniger beitragen als höhere. Unter diesen Gesichtspunkten soll das folgende Verfahren mit dem Ziel einer Bandbreiteneinsparung erläutert werden.

Unter den gegebenen Verhältnissen ist es zweifelsohne wichtig zu wissen, daß bei der menschlichen Stimme die maximale Information innerhalb des zu übertragenden Bandes bei etwa 800...1100 Hz liegt. Die tiefer oder höher liegenden Frequenzen sind zwar auch erforderlich, bringen jedoch nicht mehr Nachricht, sondern tragen lediglich dazu bei, den Klangcharakter des Sprechers erkenntlich zu machen; sie sind also energiemäßig nur zu einem geringen Teil wichtig.

Die Verständlichkeit geht bei der telefonischen Sprachübertragung (300...3400 Hz) um nur etwa 2% zurück, wenn die tiefste Frequenz 500 Hz anstatt 300 Hz beträgt, während gleichzeitig die Schallenergie um etwa 60% zurückgeht. Beschneidet man dagegen das Nachrichtenband bei etwa 1500 Hz als höchste Frequenz, dann ist die Verständlichkeitseinbuße etwa 35%, während die Energie nur um etwa 10% abnimmt. Diese Eigenschaft ist bei verschiedenen Sprachen unterschiedlich.

Diesen Erkenntnissen Rechnung tragend, wäre es zur Einsparung von Frequenzbandbreite (bei der Auswahl diskreter Frequenzen aus einem Sprachband) richtiger, mehr Signalenergie zugunsten der wichtigen tiefen und weniger für die unwichtigen hohen Bänder zu übertragen.

Die bekannten Vocoder ermöglichen es, mit sehr geringer Bandbreite auszukommen (bis 500 Hz und darunter). Allen Vocoder-Verfahren ist gemeinsam, daß am Empfangsort eine neue Sprache synthetisch aufgebaut wird, wobei nur gewisse Charakteristiken der Sprache über-

tragen werden. Dann ist zwar die Nachricht am Ende der Leitung vorhanden, die Individualität des Sprechers geht aber verloren. Am Empfangsort ist der Sprecher an seiner Stimme nicht mehr zu erkennen.

Während diese Vocoder-Verfahren dadurch zustande kommen, daß die Sprache durch eine Anzahl parallel geschalteter Filter geschickt wird und die erhaltenen Bänder nahe aneinandergerückt werden, unterscheidet sich das hier beschriebene Verfahren dadurch von jenen, daß nur je ein Filter (auf der Sende- und auf der Empfangsseite) erforderlich ist.

1. Sprachbandkompression mit Hilfe diskreter Frequenzen

Das Verfahren der Sprachübertragung mit Hilfe einer Auswahl diskreter Frequenzen (Patentschrift Nr. 953984 des Deutschen Patentamtes) erfolgt grundsätzlich über einen Speicher. Ein kürzerer Zeitabschnitt in der Größenordnung von etwa 20 ms wird zunächst in einem Speicher festgehalten. Dieser Speicher wird mehrfach nacheinander mit jeweils erhöhter Geschwindigkeit gegenüber der Aufzeichnungsgeschwindigkeit abgetastet. Die gewonnenen Bänder, deren Frequenzlage sich nach der Abtastgeschwindigkeit richtet, werden über ein relativ schmales Filter geleitet. Als Speicher kann sowohl ein magnetischer Speicher als auch ein elektronischer Speicher verwendet werden, wobei dem elektronischen Speicher der Vorzug zu geben ist²⁾.

1.1. Umsetzung der Nachricht in eine andere Frequenzlage über einen elektronischen Speicher

In einem Verfahren zur Nachrichtenübermittlung nach Patentschrift Nr. 940656 des Deutschen Patentamtes wurde bereits ein elektronischer Speicherschirm als Nachrichtenspeicher beschrieben. Solche Speicher ermöglichen immer das Einsparen von Bandfiltern. Während im vorgenannten Fall sämtliche Elektronenstrahlen mit der gleichen Geschwindigkeit über den Speicher streichen, wird in dem nachstehend beschriebenen Verfahren nur ein Elektronenstrahl benutzt, der aber mit unterschiedlicher Geschwindigkeit den Speicherschirm mehrfach überstreicht. Unter Weglassung sämtlicher zum Verständnis nicht erforderlicher Teile ist im Bild 1 das Prinzip des Verfahrens dargestellt.

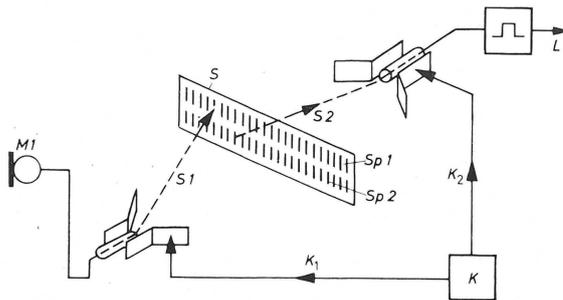
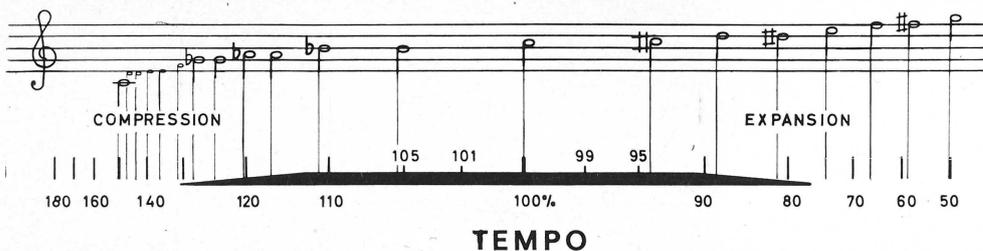


Bild 1. Schematische Darstellung eines Verfahrens zur Sprachbandkompression mittels Umsetzung der Information in eine andere Frequenzlage mit Hilfe eines elektronischen Speichers

²⁾ Springer, A. M.: Beitrag zur Übertragung oder Aufzeichnung eines breiten Nachrichtenbandes. Elektron. Rdsch. Bd. 16 (1962) Nr. 6, S. 259—260 Patentschrift Nr. 1003809 des Deutschen Patentamtes

TONLAGEN-REGLER ohne Geschwindigkeitsänderung



oder

TEMPO REGELUNG

ohne

Tonlagenänderung

für Magnetton-Schallaufnahmen mit 38 cm/sec

Eltro

GMBH & CO., GESELLSCHAFT FÜR STRALUNGSTECHNIK, HEIDELBERG

Die vom Mikrophon *M 1* kommende Nachricht mit einer Bandbreite von beispielsweise 300...3000 Hz wird nach Verstärkung einem Elektronenstrahl aufmoduliert. Der Elektronenstrahl *S 1* bewegt sich mit konstanter Geschwindigkeit *v* über einem elektronischen Speicher *S* und lädt die einzelnen Mosaikkondensatoren des Speichers auf — in der Zeit τ_1 beispielsweise die obere Spur *Sp 1*, in der Zeit τ_2 die untere Spur *Sp 2* des Speicherschirms *S*, dann wieder die obere Spur usw. Ein zweiter Elektronenstrahl, der Abtaststrahl *S 2*, tastet (ebenfalls sägezahnförmig) die Spuren ab, jedoch nach folgender Art: Nachdem die Spur *Sp 1* durch den Elektronenstrahl *S 1* moduliert worden ist, bewegt sich der Abtaststrahl *S 2* beispielsweise mit der dreifachen Geschwindigkeit, mit der die Aufzeichnung stattgefunden hat (also mit $3v$) über die Speicherspur *Sp 1*. Dazu verbraucht der Abtaststrahl nur ein Drittel der Aufspeicherzeit (Bild 2). Hierauf bewegt sich der Strahl *S 2* mit der sechsfachen Geschwindigkeit ($6v$) über die gleiche Abtastspur in einem Sechstel der Zeit, dann mit der achtfachen Geschwindigkeit in einem Achtel der Zeit, mit der zehnfachen Geschwindigkeit, mit der elffachen, dreizehnfachen, achtzehnfachen und schließlich mit der dreißigfachen Geschwindigkeit.

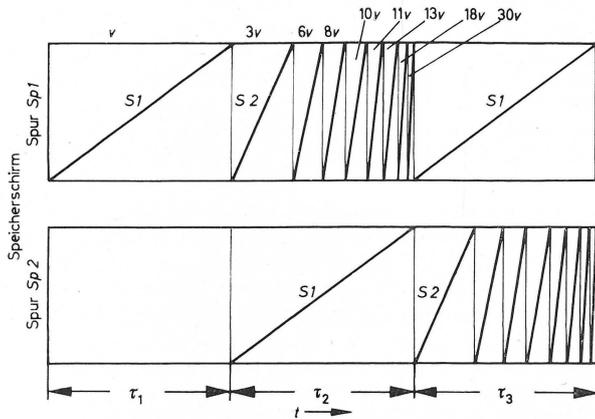


Bild 2. Zeitgeschwindigkeitsdiagramm der Elektronenstrahlen *S 1* und *S 2*; die im Beispiel angeführte achtfache Abtastung (Spur *Sp 1* in der Zeit τ_2 , Spur *Sp 2* in der Zeit τ_3) dauert ebenso lange wie der Aufzeichnungsvorgang (Spur *Sp 1* in der Zeit τ_1 , Spur *Sp 2* in der Zeit τ_2)

Das im Bild 1 angedeutete Kippgerät *K* steuert die Kippfrequenzen K_1 und K_2 . Über einen Bandpaß gelangt die Nachricht auf die Übertragungsleitung *L*.

Die Summe der einzelnen Abtastzeiten ergibt nach Gl. (1) im angeführten Beispiel der achtfachen Abtastung 98% der Aufspeicherzeit. Die restlichen 2% werden für die Rückführung des Strahls benötigt.

$$\sum \frac{1}{n} = \left(\frac{1}{3} + \frac{1}{6} + \frac{1}{8} + \frac{1}{10} + \frac{1}{11} + \frac{1}{13} + \frac{1}{18} + \frac{1}{30} \right) = 0,98 \quad (1)$$

Die nacheinander erfolgende Abtastung der gleichen Aufzeichnung in acht verschiedenen Zeiten ergibt acht Bänder. Die einzelnen Bänder liegen in entsprechend angehobenen Frequenzlagen. Im Bild 3 sind

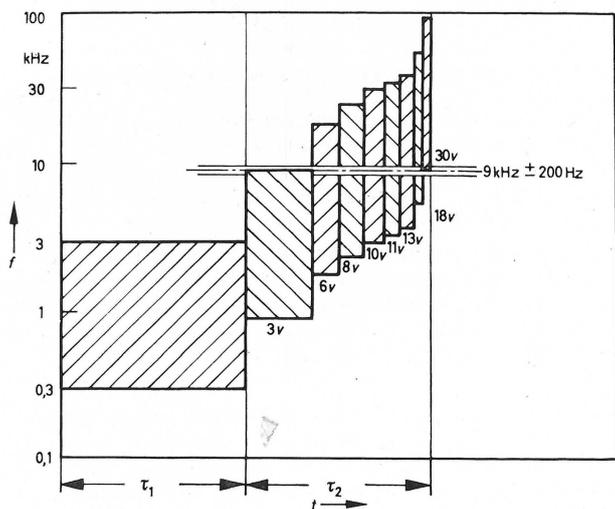


Bild 3. Durch acht verschiedene Abtastgeschwindigkeiten erreichte Frequenzlage der Nachricht

auf Grund der verschiedenen Abtastgeschwindigkeiten $3v, 6v, 8v, 10v, 11v, 13v, 18v, 30v$ die einzelnen Lagen schematisch angedeutet. Jedes einzelne abgetastete Band enthält die volle Information und erlaubt eine Beschneidung des Signals bei der Übertragung. Die gegenüber der ursprünglichen Aufzeichnung mit verkürzten Zeiten abgetasteten Bänder werden über ein Bandfilter geführt, so daß aus jedem einzelnen Band nur ein schmaler Bereich übrigbleibt.

Die in der Zeit τ_1 gesendete Information mit einer angenommenen Bandbreite von $0,3...3$ kHz ist im Bild 3 in der ursprünglichen Frequenzlage im schraffierten Rechteck über τ_1 enthalten. Im ersten Abschnitt über τ_2 ist die gleiche Information wieder enthalten, da die gesamte Speicherspur *Sp 1* vom Strahl *S 1* abgetastet wurde. Da dies aber mit der dreifachen Geschwindigkeit erfolgte, hat sich der ursprüngliche Frequenzbereich von $0,3...3$ kHz um den Faktor 3 nach oben verschoben und liegt nun im Bereich von $0,9...9$ kHz. Der nächste Abschnitt, der mit der sechsfachen Geschwindigkeit $6v$ vom gleichen Speicher abgetastet wurde, enthält wiederum die volle Nachricht, jedoch in der neuen Frequenzlage $1,8...18$ kHz. Das mit der achtfachen Geschwindigkeit $8v$ abgetastete dritte Band liegt nun im Bereich von $2,4...24$ kHz, das zehnte Band liegt dann bei $3...30$ kHz, das elfte bei $3,3...33$ kHz, das dreizehnte bei $3,9...39$ kHz, das achtzehnte bei $5,4...54$ kHz, und das letzte Band, das nur ein Dreißigstel der Aufspeicherzeit von τ_1 für die Abtastung benötigt, liegt im Bereich von $9...90$ kHz.

Die Information, die mit dem Abtaststrahl abgenommen wurde, ist demnach in den Bereich von $0,9...90$ kHz transponiert. Wird diese Information nun, wie im Bild 3 angedeutet, über ein schmales Filter im 9-kHz-Bereich geführt, der nur eine Bandbreite von ± 200 Hz hat, so werden schmale Stücke aus den einzelnen Bändern herausgeschnitten. Der Informationsinhalt entspricht der im Bild 4 ge-

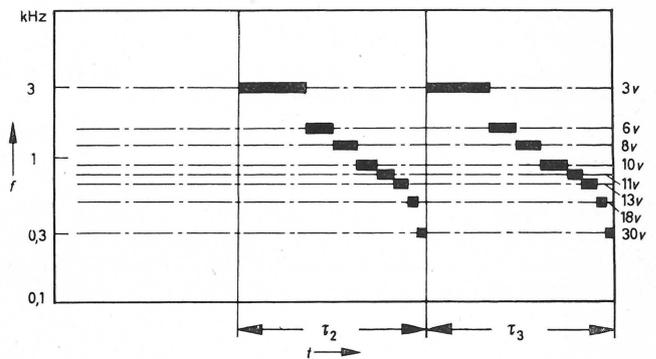


Bild 4. Entsprechend den verschiedenen Abtastgeschwindigkeiten aufgespaltenes Frequenzband der Nachricht

zeichneten Information. Die einzelnen Teilstücke der herausgeschnittenen diskreten Frequenzen unterscheiden sich lediglich in der Amplitude untereinander, da die einzelnen schmalen Bänder unterschiedliche Energieanteile haben.

Im Bild 5 sind die auf den Frequenzbereich der Originalinformation bezogenen Bandbreiten und diskreten Frequenzen der einzelnen Bänder angegeben.

Ein Abhören einer solchen Nachricht ist praktisch ausgeschlossen, denn zunächst liegt die ganze Information in einem 9-kHz-Kanal, der nichts von der Nachricht mehr erkennen läßt. Dieses im 9-kHz-Band liegende Signal läßt sich durch verschiedene Mittel übertragen. Eines davon ist besonders zweckmäßig: Die Nachricht wird durch einen Träger zum Beispiel auf etwa 600 Hz herabgesetzt. Dieser Träger muß selbstverständlich an der Empfangsstelle der Information wieder zugesetzt werden.

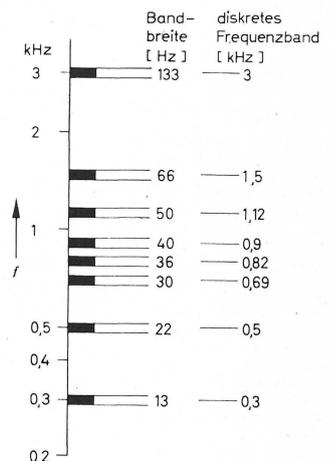


Bild 5. Bandbreite und diskrete Frequenzen des Beispiels

Am Empfangsort spielt sich der bisher beschriebene Vorgang in umgekehrter Weise ab, jedoch ist ein Filter grundsätzlich nicht mehr erforderlich. Das Empfangssignal wird mit einem Elektronenstrahl mit der gleichen Geschwindigkeit, mit der es sendeseitig abgetastet wurde,

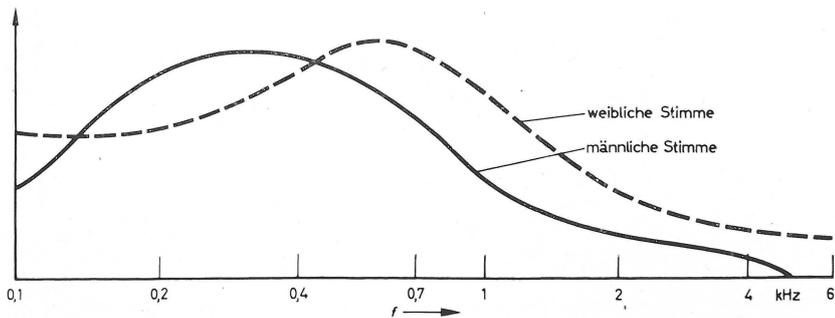


Bild 6. Die zur Übertragung dienenden Bänder lassen sich nach dem Energieinhalt der Nachricht verschieben, beispielsweise männliche und weibliche Stimmen

auf einer Spur des Speicherschirms aufgezeichnet und mit einem zweiten, sich langsam bewegenden Strahl (entsprechend dem Strahl *S I* bei der Aufzeichnung) abgetastet.

Während auf der Sendeseite der Aufzeichnungsstrahl bereits während der Aufzeichnung die darunterliegende Spur löscht, löscht auf der Empfangsseite der Abtaststrahl die Spur.

In Abwandlung dieser Lösung kann auch der Aufzeichnungsstrahl der Sendeseite durch einen fächerartigen Strahl, der eine breite Tonspur aufzeichnet, ersetzt werden. Der Abtaststrahl, der nun eine punktförmige Bündelung hat, bewegt sich bei der mehrfachen Abtastung langsam nach oben oder unten, so daß er nur das noch nicht entladene Mosaik des Speichers abtastet. Auf der Empfangsseite zeichnet der Speicherstrahl bei der mehrfachen Aufzeichnung in mehreren Linien übereinander auf, und der Abtaststrahl überstreicht nun fächerartig sämtliche Linien gleichzeitig. Die Nachricht befindet sich dann in der ursprünglichen Frequenzlage.

1.1.1. Eigenschaften des Verfahrens

Das Verfahren, das unter dem deutschen Patent Nr. 953984 bereits seit 1953 geschützt ist, hat folgende Eigenschaften:

1. Ein Abhören auf dem Übertragungswege ist ohne gleiche Einrichtung unmöglich.
2. Sämtliche Frequenzen, die an der Übertragung teilnehmen, haben bei diesem Verfahren die gleichen Laufzeiten. Das bedingt, daß die sonst üblichen Laufzeitkompensationen wegfallen können. Damit erhöht sich auch der Abstand der erforderlichen Verstärker. Die Verstärker brauchen nur nach der Leitungsdämpfung ausgelegt zu werden.
3. Die Nachricht kann in der Bandbreite beliebig verändert werden. Durch die Wahl der Kippfrequenz kann auch die Anzahl der einzelnen diskreten Bänder erhöht werden.
4. Die Bandbreite sämtlicher diskreter Bänder läßt sich mit Hilfe eines einzigen Filters verändern.
5. Die einzelnen diskreten Bänder können je nach dem Energieanteil der Sprache untereinander verschoben werden (Bild 6).

1.2. Umsetzung der Nachricht in eine andere Frequenzlage über einen magnetischen Speicher

Das beschriebene Übertragungsverfahren läßt sich anstatt mit elektronischen Mitteln auch mit magnetischen Speichern aufbauen. Es besteht beispielsweise die Möglichkeit, den gleichen Effekt auch magnetisch dadurch zu erreichen, daß mehrere umlaufende Magnettonköpfe mit unterschiedlicher Geschwindigkeit einen endlosen Tonträger, der als Speicher dient, abtasten, wobei die Nachricht in verschiedene Frequenzlagen gebracht wird. Die endlose Tonbandschleife nach Bild 7 enthält außer den normalen Sprech-, Hör- und Löschköpfen (*Spr*, *Hö*, *Lö*) zum Aufsprechen der Nachricht beispielsweise die angedeuteten vier mit verschiedener Geschwindigkeit umlaufenden

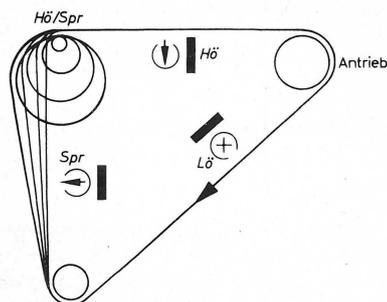


Bild 7. Umsetzung einer Nachricht mit Hilfe eines Tonbandspeichers in verschiedene Frequenzlagen durch vier mit verschiedenen Geschwindigkeiten umlaufende Magnettonköpfe

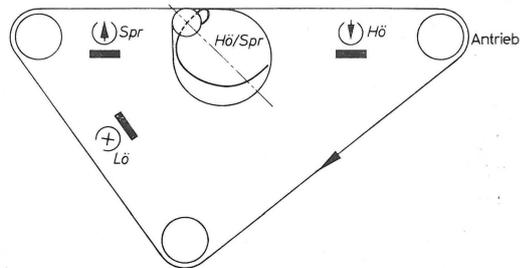


Bild 8. Umsetzung einer Nachricht mit Hilfe eines einzigen umlaufenden kegelartigen Magnettonkopfes mit spiralförmig angeordnetem Magnetpalt

den Magnettonköpfe *Hö/Spr*. Sendeseitig arbeiten diese umlaufenden Köpfe als Hörköpfe, empfangsseitig als Sprechköpfe.

Ein weiterer Weg, mit magnetischen Speichern eine Frequenzbandverschiebung durchzuführen, läßt sich in ähnlicher Art auch mit einem einzigen umlaufenden Magnettonkopf erreichen, dessen Magnettonspalt als Spirale an der Oberfläche eines Kegels verläuft (Bild 8). Solche magnetischen Tonköpfe sind labormäßig auch gebaut worden. Der Kegel des umlaufenden Kopfes *Hö/Spr* dreht sich um die eingezeichnete Achse, wobei der Spalt punktweise den Tonträger berührt und dabei den gesamten Frequenzbereich mit unterschiedlicher Geschwindigkeit sägezahnförmig abtastet oder aufspricht (Bild 9).

Im Bild 9 ist das Frequenzzeitdiagramm nach dem im Bild 8 gezeigten Verfahren dargestellt. Der Sprechkopf dreht sich in der gezeichneten Lage mit konstanter Geschwindigkeit. Dabei wird das aufzuzeichnende Signal am kleinen Durchmesser mit geringerer, am großen Durchmesser mit höherer Geschwindigkeit aufgezeichnet. Je nach der Drehrichtung enthalten diese parallelogrammartigen Frequenzbänder den ursprünglichen Nachrichteninhalt und liegen aufwärtsgehend oder abwärtsgehend. Ein ausgefiltertes Frequenzband, das mit dem gleichartigen synchronlaufenden Kopf auf der Empfangsseite aufgezeichnet wird, gibt dann die im Bild 10 angegebenen Teilbänder

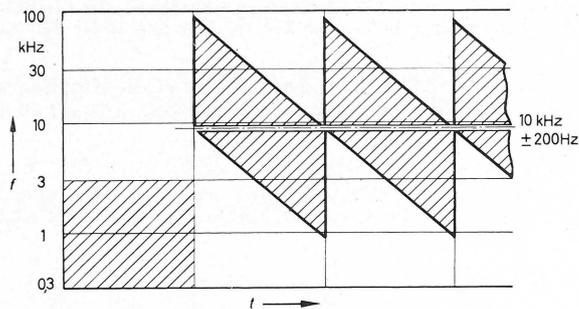


Bild 9. Mit der Anordnung nach Bild 8 verwandelte Nachricht in ihrer Frequenzlage

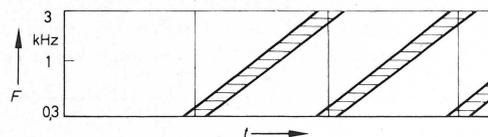


Bild 10. Auf die Empfangsseite zurückverwandelte Nachricht

wieder. Wie sich diese Nachricht letztlich anhören wird, muß dem Versuch überlassen bleiben. Zweifelsohne ist dabei der technische Aufwand der Mechanik größer und die Lebensdauer kürzer als bei der elektronischen Lösung. Außerdem erhöht sich bei allen magnetischen Speichern die unumgängliche Speicherzeit.

In der Nachrichtentechnik, die sich immer mehr der Elektronik bedient, wird man mechanische Bewegungen, wie sie zur Aufteilung des Frequenzbandes mit dem zuletzt beschriebenen Verfahren gezeigt wurden, wohl kaum verwenden. Es sollte nur die Möglichkeit einer anderen Lösung dargestellt werden.

*

Im Zeitalter der außerordentlich großen Entfernungen zwischen Sender und Empfänger und der gleichzeitig von sämtlichen Postanstalten der Welt verlangten Geheimhaltung bedeutet das beschriebene Sprachübertragungsverfahren der Nachrichtenübermittlung ohne Zweifel eine weitere Möglichkeit.

(eingegangen am 16. Mai 1963)