



SCHWEIZERISCHE EIDGENOSSENSCHAFT
BUNDESAMT FÜR GEISTIGES EIGENTUM

(51) Int. Cl.³: G 10 L

1/10

Erfindungspatent für die Schweiz und Liechtenstein

Schweizerisch-liechtensteinischer Patentschutzvertrag vom 22. Dezember 1978



(12) PATENTSCHRIFT A5

(11)

625 900

(21) Gesuchsnummer: 9970/77

(22) Anmeldungsdatum: 15.08.1977

(30) Priorität(en): 16.08.1976 US 714495

(24) Patent erteilt: 15.10.1981

(45) Patentschrift
veröffentlicht: 15.10.1981

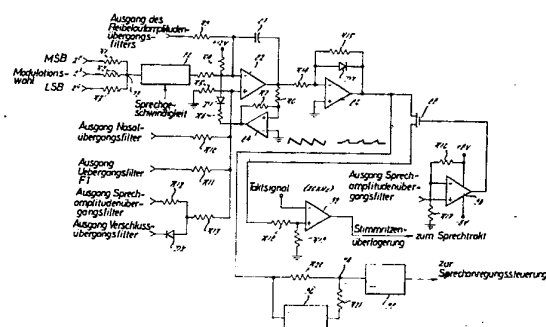
(73) Inhaber:
Federal Screw Works, Detroit/MI (US)

(72) Erfinder:
Mark Vincent Dorais, Detroit/MI (US)

(74) Vertreter:
E. Blum & Co., Zürich

(54) Sprachsynthetisator.

(57) Der Sprachsynthetisator enthält Steuerschaltkreise, die auf Eingangsbefehlsworte ansprechen, um ein vernehmbares Ruhephonem in das Sprechmuster einzuführen, die Dauer jedes erzeugten Phonem zu variieren und die Gesamtgeschwindigkeit und das Volumen der erzeugten Sprache zu ändern. Zusätzlich ist eine Modulationseinrichtung (80) zur Erzeugung einzelner Phoneme und einer Stimmritzenwellenform, welche der menschlichen Stimmritze besser entspricht, sowie eine Überlagerungsschaltung vorhanden, die im Stimmtrakt eine Resonanzunterdrückung bewirkt, um bei Öffnung der Stimmritze einen Dämpfungseffekt zu erzielen und einen höheren menschlichen Stimmenergieinhalt bei höheren Frequenzen zur Verbesserung der erzeugten Sprachqualität zu erzeugen.



PATENTANSPRÜCHE

1. Sprachsynthesator, gekennzeichnet durch eine Eingabe-einrichtung (12, 14), die auf Eingabedaten, welche eine gewünschte Phonemfolge bestimmen, anspricht, um Steuersignale zu erzeugen, welche die bestimmte Phonemfolge bestimmenden Parameter darstellen, durch eine Sprechquelle (54) zur Erzeugung eines stimmhaften Anregungssignals mit einer Grundfrequenz, durch eine Ausgabevorrichtung (60, 62, 66, 68, 70) mit einem Stimmtrakt (60), der auf das stimmhafte Anregungssignal und die Steuersignale anspricht, um die Phonemfolge zu erzeugen, und durch eine an die Sprechquelle (54) angeschlossene Modulationseinrichtung (56) zur automatischen Modulation der Grundfrequenz des stimmhaften Anregungssignals in Übereinstimmung mit bestimmten Steuersignalen, die durch die Eingabeeinrichtung erzeugt werden.

2. Sprachsynthesator nach Anspruch 1, dadurch gekennzeichnet, dass die Ausgabevorrichtung einen Stimmtrakt (60), drei variable Resonanzfilter (F1-F3), die jeweils durch eines der Steuersignale abstimmbar sind, um die ersten drei Formanten in den Frequenzspektren zu erzeugen und einen vierten abstimmbaren Resonanzfilter (F4) enthält, der durch ein einen der drei Resonanzfilter abstimmdes Steuersignal abstimmbar ist, um einen vierten Formanten in den Frequenzspektren zu erzeugen.

3. Sprachsynthesator nach Anspruch 1, gekennzeichnet durch eine Reibelautquelle (64) zur Erzeugung eines stimmlosen Anregungssignals, wobei die durch die Eingabeeinrichtung (12, 14) erzeugten Steuersignale ein erstes Steuersignal zur Steuerung der Amplitude des stimmhaften Anregungssignals und ein zweites Steuersignal zur Steuerung der Amplitude des stimmlosen Anregungssignals enthalten und wobei der Stimmtrakt (60) auf die stimmhaften und stimmlosen Anregungssignale und die Steuersignale anspricht, um einen Audioausgang, bestehend aus einer die menschliche Sprache nachahmenden Phonemfolge, zu erzeugen, und durch eine Amplitudensteuer-einrichtung (36, 40, 50) zur Änderung der Gesamtamplitude des Audioausganges durch Modulation eines vorausgewählten Signalmerkmals des ersten und zweiten Steuersignals.

4. Sprachsynthesator nach Anspruch 3, dadurch gekennzeichnet, dass die Amplitudensteuereinrichtung (36, 40, 50) auf bestimmte Eingabedaten anspricht, um die Gesamtamplitude des Audioausganges bei Aufrechterhaltung der relativen Amplitudenänderungen im stimmhaften und stimmlosen Anregungssignal zu ändern, welche Amplitudenänderungen von Phonem zu Phonem durch die ersten bzw. zweiten Steuersignale durch fortlaufende Modulation des vorausgewählten Signalmerkmals der ersten und zweiten Steuersignale um einen bestimmten Prozentsatz auftreten.

5. Sprachsynthesator nach Anspruch 1, gekennzeichnet durch eine Takteinrichtung (32), die auf eines der Steuersignale anspricht, um ein Taktsignal zu erzeugen, das den Erzeugungszeitraum für jedes Phonem bestimmt, und durch eine Geschwindigkeitsteuereinrichtung (28, 30, 50), die auf die Eingabedaten anspricht, um die Phonemtaktung durch Erzeugen eines Sprechgeschwindigkeitssignals in Abhängigkeit von den Eingabedaten, die der Takteinrichtung (32) zur Änderung des Taktsignals zugeleitet werden, zu ändern.

6. Sprachsynthesator nach Anspruch 1, gekennzeichnet durch eine Reibelautquelle (64) zur Erzeugung eines stimmlosen Anregungssignals, und durch eine auf die Eingabedaten ansprechende Schaltung (36, 40, 50), die auf den Stimmtrakt einwirkt, um durch Unterdrückung der Anregung des Stimmtraktes (60) durch das stimmhafte und stimmlose Anregungssignal ein Ruhephonem zu erzeugen.

7. Sprachsynthesator nach Anspruch 1, dadurch gekennzeichnet, dass eine Einrichtung vorgesehen ist, um einen Frequenzformanten mit einer bezüglich des Sprachfrequenzbandes relativ hohen Frequenz zum stimmhaften Anregungssignal

zur Erhöhung des Energieinhalts des stimmhaften Anregungssignals zu addieren.

8. Sprachsynthesator nach Anspruch 7, dadurch gekennzeichnet, dass der Stimmtrakt eine Mehrzahl von Resonanzfiltern (F1-F5) aufweist, wobei mindestens ein Resonanzfilter mit einer höheren Frequenz als der zum stimmhaften Anregungssignal addierte Frequenzformant schwingt.

9. Sprachsynthesator nach Anspruch 6, dadurch gekennzeichnet, dass der Stimmtrakt dazu bestimmt ist, entsprechend den Steuersignalen das Aussprechmuster der aufeinanderfolgenden Phoneme, die durch die Eingabedaten während der Erzeugung des Ruhephonems identifiziert werden, zu bilden.

Die vorliegende Erfindung bezieht sich auf einen Sprachsynthesator zum phonetischen Synthetisieren der menschlichen Sprache.

Die Entwicklung auf diesem Gebiet führte bis vor kurzem zur Herstellung von sehr komplizierten und teuren Einrichtungen, welche einen sehr unnatürlichen Klang in der Sprache erzeugten. Dies ist hauptsächlich auf die Tatsache zurückzuführen, dass diese erste Generation von Synthesatoren ohne eine genügende Entwicklung hergestellt wurden, mit dem Versuch, einen Synthesator zu schaffen, der in der Lage sein sollte, jede bekannte Funktion der menschlichen Sprache auszuführen. Die so erhaltenen Einrichtungen waren folglich nur in der Lage, einige Funktionen sicher durchzuführen.

Typisch für den Aufbau dieser früheren Sprachsynthesatoren war die Behandlung der entsprechenden Übergangsperioden zwischen den Phonemen. Durch die Erkennung der Wichtigkeit der Übergangsperioden in der menschlichen Sprache ergab sich bei einigen Einrichtungen im wesentlichen das Bestreben, verschiedene Übergangswellenformen zu erzeugen, um die menschliche Aussprache zwischen den Lautzuständen zu simulieren. Die sehr komplexen Schaltungen, die erforderlich sind, um die Erzeugung dieser Wellenformen in einwandfrei phonetischer Sprache zu analysieren, zu steuern und zusammenzufassen, machten die Einrichtungen für den gewöhnlichen Gebrauch sehr unpraktikabel.

Die letzten Entwicklungen auf diesem Gebiet erbrachten die Tatsache, dass die genaue Nachbildung des menschlichen Sprachsystems ein unerreichbares Ziel ist und führte dazu, eine Annäherung an das menschliche Sprechsystem zu finden, welche eine akzeptierbar klingende Sprache erzeugt. Ohne die Wichtigkeit der Zwischenlautübergänge herabzusetzen, war das Hauptergebnis dieser Entwicklung der Wechsel von dem sehr komplexen System der Zwischenlautübergänge zu einem einfachen Lösungsweg, bei dem relativ langsam arbeitende Filter verwendet werden, die die abrupten Änderungen in den Steuerungsparametern, welche die Zustände zwischen den einzelnen Phonemen bestimmen, glätten.

Demzufolge ist es das Hauptziel der vorliegenden Erfindung einen Sprachsynthesator zu schaffen, der nicht nur relativ unkompliziert und billig ist, sondern auch in der Lage ist, eine auffallend natürlich klingende Sprache zu erzeugen.

Ein weiteres Ziel der Erfindung ist es, einen Sprachsynthesator zu schaffen, der diese natürlich klingende Sprache erzeugt, wodurch diese Einrichtung sich insbesondere zur Verwendung mit einem Digitalrechner als ein Text-Sprach-Wandler eignet.

Das bevorzugte Ausführungsbeispiel der vorliegenden Erfindung enthält eine Einrichtung, die digitalisierte Daten, wie das Ausgangssignal eines Computers oder anderer Digitalvorrichtungen, durch Erzeugung und Integration der Phoneme der Sprache in eine elektronisch synthetisierte menschliche

Sprache umwandelt. Das digitale Basisbefehlswort, welches den Sprachsynthesator treibt, weist mit Vorteil zwölf Bits auf. Sieben dieser Bits sind der Phonemauswahl zugeteilt, um ein bestimmtes Phonem, eine Pause oder Steuerfunktion zu bestimmen. Somit erhält man ein Maximum von 2^7 oder 128 unterschiedliche Befehle. Diese über das zur Erzeugung der Grundphonemklänge hinaus erweiterte Kapazität ermöglicht es der vorliegenden Vorrichtung, eine grössere Variation von Allophonen, welche die leicht geänderten Grundphoneme darstellen, nachzubilden, um sie leichter in die Sprachschwankungen zu integrieren. Z. B. der Laut «æ» im englischen Wort «happen» unterscheidet sich von dem im englischen Wort «bat». Der Anfangslaut «k» im englischen Wort «kick» unterscheidet sich von dem im englischen Wort «quit». Diese erweiterte Kapazität erlaubt es der vorliegenden Vorrichtung verschiedene Befehle zur Erzeugung von gleichen Lauten in verschiedene Sprachen zur Verfügung zu stellen.

Drei der zwölf Datenbits im Eingabebefehlswort werden zur Steuerung der Sprachmodulation verwendet. Dies ergibt 2^3 oder acht Modulationspegel pro Phonem, was der Vorrichtung die Fähigkeit verleiht, die glatten und feinen Bewegungen in der Toncharakteristik der menschlichen Sprache nachzubilden. Die verbleibenden zwei Datenbits in jedem Eingabebefehlswort werden zur Änderung der Geschwindigkeit der Phonemerzeugung verwendet. Es sind vier mögliche Zeitintervalle für jedes erzeugte Phonem vorgesehen, um die Phoneme besser an die Zeitdauer anzupassen.

Die sieben Bits, welche das einzelne Phonem bestimmen, werden einer Eingabeschaltung zugeführt, die eine Mehrzahl vorbestimmter Steuersignalparameter elektronisch erzeugen, die das ausgewählte Phonem bestimmen. Die durch die Eingabeschaltung erzeugten Steuersignale werden mit Vorteil in Form von reihenförmig, binär-gewichteten Rechtecksignalen erzeugt, deren Mittelwerte gleich den sie darstellenden Analogsignalen sind. Durch Erzeugung von Analogsignalen darstellenden Digitalsignalen werden bei der vorliegenden Vorrichtung keine komplizierten Schaltungen zur genauen Steuerung der Analogsignale verwendet.

Die Steuersignalparameter aus der Eingabeschaltung durchlaufen zuerst eine Reihe von langsam arbeitenden Übergangsfiltren, welche die abrupten Amplitudenänderungen im Signal glätten. Danach werden die Steuersignale zu verschiedenen Schaltungen zur Steuerung der dynamischen Artikulation geleitet, welche die Parameter kombinieren und verarbeiten, um Schallerregung und Tontraktsteuersignale zu erzeugen, die analog zu den Muskelbefehlen aus dem Gehirn an den Tontrakt, die Stimmritze, die Zunge und Mund der menschlichen Sprechorgane sind.

Die Vorrichtung enthält ferner Vokal- und Reiblautanregungsquellen, welche die Anregungssignale empfangen, die die verschiedenen Signaleigenschaften der stimmhaften und stimmlosen Basissignalmengen in der menschlichen Sprache bestimmen. Die Vokalanregungsquelle erzeugt eine Stimmritzenwellenform, welche die Stimmritze, wie sie im menschlichen Kontakt schwingt, nachahmt. Die Reibelautquelle simuliert den Klang einer durch eine eingeschränkte Öffnung strömenden Luft, wie er beim Ansprechen der Laute s, f und h auftritt.

Die Vokal- und Reiblautanregungssignale sowie die Sprechtraktsteuersignale werden an eine Reihe kaskadenförmig geschalteter frequenzabhängige Filter angelegt, welche die Vielzahl der Resonanzkammern im menschlichen Tontrakt simulieren. Die Steuersignale stellen die Eigenresonanzen des Filters ein, um ein Audiosignal mit dem gewünschten Frequenzspektrum zu erzeugen.

Die zwei Taktbits im ursprünglichen Eingabebefehlswort werden zu einem Einschaltdauerakt signal umgewandelt, welches dem Phonemtaktgeber, der das Zeitintervall des erzeug-

ten einzelnen Phonems bestimmt, zugeleitet wird. Die drei verbleibenden Modulationsbits im Eingabebefehlswort werden zur Erzeugung eines analogen Modulationssteuersignals verwendet, das der Vokalanregungsquelle zugeleitet wird, um die «Teilung» oder Frequenz der Stimmritzenwellenform zu bestimmen.

Ferner ist eine Schaltung vorhanden, die die Modulationshöhen von verschiedenen Phonemen entsprechend dem bestimmten Parametersteuersignal automatisch wechselt.

Daraus ergibt sich, dass die durch die vorliegende Vorrichtung erzeugte Stimme weniger monoton ist und einen natürlicheren Klang hat als die bisher bekannten Vorrichtungen, insbesondere wenn sich das manuelle Programmieren unpraktisch erweist oder nicht angewendet wird.

Es ist möglich, ausserdem eine neue Stimmritzenwellenform anzuwenden, welche die Tätigkeiten der menschlichen Stimmritze genauer simuliert. Die neue Wellenform kann aus einer abgestumpften Sägezahnwelle bestehen, welche sowohl ungerade und gerade Harmonische erzeugt. Ausserdem kann in dieser Wellenform zusätzlich ein Hochfrequenzformant enthalten sein, der die Spektralenergie der Wellenform bei hohen Frequenzen steigert. Durch die gesteigerte Energie bei hohen Frequenzen kann die Amplitude der niederen Formanten vergrössert werden.

Der Vokaltrakt kann ebenfalls verbessert werden, und zwar durch Verstellung des Resonanzfilters vierter Ordnung im Vokaltrakt. Dies ist insbesondere bedeutsam, weil dies ohne die Erfordernis der Erzeugung von zusätzlichen Steuerparametern, welche die Vorrichtung nur komplizieren würden, ausgeführt wird. Dieser Filter wird von dem gleichen Steuersignal verstellt, das die Stelle des dritten Resonanzpoles bestimmt.

Im Vokaltrakt können die Tonresonanzen überlagert werden, um die reduzierte Dämpfung, die im menschlichen Tontrakt bei offener Stimmritze reflektiert wird, zu simulieren.

Dazu ist eine Schaltung zur Erzeugung eines Rechtecksignals mit veränderlicher Impulsweite vorgesehen, dessen Einschaltdauer proportional zur Amplitude der Stimmritzenwellenform ist. Dieses Überlagerungssignal wird dann einer Reihe von Analoggattern zugeleitet, die an den Bandpassteilen der ersten drei Resonanzfilter im Tontrakt angeschlossen sind. Bei offener Stimmritze wird durch die Vergrösserung der Bandweiten der Resonanzfilter eine Dämpfung der Resonanz erzielt, wenn sich die Amplitude der Stimmritzenwellenform vergrössert.

Schliesslich kann noch eine Kennzeichenbefehlskodier- und Steuerschaltung vorgesehen werden, welche den Programmierer in die Lage versetzt, die gesamte Lautstärke und Sprechgeschwindigkeit des Audioausgangs zu ändern. Mit dieser Schaltung kann auch ein Ruhephonem in die Sprachform eingeführt werden, wobei das Ruhephonem in der gleichen Weise wie ein stimmhaftes Phonem ausgesprochen werden kann, um die Natürlichkeit der erzeugten Sprache zu verbessern. Das Ruhephonem kann primär zur Verwendung mit bestimmten Phonemen verwendet werden, die natürlicher klingen, wenn ihre Ausdrucksform vorher gebildet wird oder ein kurzer Zeitraum nach dem Anlegen der Anregungsenergie an den Vokaltrakt erhalten wird.

Die Kennzeichenschaltung kann durch einen spezifischen 7-Bit-Phonemcode erregt werden, welcher den Kennzeichenbefehl von den anderen Phonembefehlen unterscheidet. Die verbleibenden fünf Bits im Kennzeichenbefehlswort werden dann zur Auswahl der Klanghöhe und Sprachgeschwindigkeit und zur Anzeige, ob die nachfolgende Phonemperiode eine Ruheperiode ist, verwendet. Das Kennzeichenbefehlsphonem verbraucht ein sehr kurzes Zeitintervall, so dass das normale phonetische Verarbeiten einer Nachricht unmerklich gewechselt wird. Dies kann durch Sperren der Kennzeicheninformation erreicht werden. Dem Synthesator wird befohlen, sofort mit dem nächsten Phonem fortzufahren.

Im folgenden ist ein Ausführungsbeispiel des Erfindungsgegenstandes näher erläutert. Es zeigen:

Fig. 1a und 1b ein Blockschema eines Ausführungsbeispiels eines Sprachsynthetisators gemäss der vorliegenden Erfindung,

Fig. 2 ein Schaltschema eines Teils des in Fig. 1 dargestellten Synthetisators,

Fig. 3a ein Diagramm, welches das Verhältnis zwischen der Grundfrequenz der Stimmritzenwellenform und der Verschiebung des ersten Resonanzpoles über einen Vokalphonembereich aufzeigt,

Fig. 3b ein Diagramm, welches den Wechsel in der Grundfrequenz der Stimmritzenwellenform über einen Bereich konstanter Phonemumgebung aufzeigt,

Fig. 4 ein Diagramm der Spektralenergie der Stimmritzenwellenform vor und nach der Hochpolkompensation,

Fig. 5 ein Schaltschema des dritten und vierten Resonanzfilters im Vokaltrakt des Synthetisators in Fig. 1,

Fig. 6 ein Schaltschema der Kennzeichenbefehlkodier- und Steuerschaltung,

Fig. 7 ein Diagramm, welches die Modulation der Amplitudensteuersignale aufzeigt, und zwar wie sie durch die Kennzeichenbefehlkodier- und Steuerschaltung in Fig. 6 erzeugt werden, und

Fig. 8 ein Diagramm, welches die Taktung der verschiedenen Taktsignale in der Kennzeichenbefehlkodier- und Steuerschaltung aufzeigt.

Die Figuren 1a und 1b zeigen ein Blockschaltbild eines Ausführungsbeispiels eines Sprachsynthetisators. Dieser Sprachsynthetisator wird mit Vorteil durch ein Zwölf-Bit-Eingabebefehlswort betrieben. Sieben der zwölf Eingabebits werden zur Phonemauswahl verwendet und werden einem Paar von Festspeichern 12 und 14 zugeleitet. Für jedes der 128 möglichen Phoneme, das durch die sieben Bits für die Phonemauswahl identifiziert werden kann, sind in den Festspeichern 12 und 14 sechzehn unterschiedliche Parameter gespeichert, die jedes Phonem elektronisch bestimmen. Jeder Parameter erfordert vier Auslösungsbits, um ein reihenförmiges binär-gewichtetes digitales Steuersignal zu erzeugen. Der gesamte Bedarf an Festspeicherbits beträgt deshalb $16 \cdot 4 \cdot 128 = 8192$ Bits. Dieser Bedarf kann durch die Verwendung von parallelgeschalteten Festspeichern gedeckt werden. Der in Fig. 1 gezeigte Synthetisator enthält zwei Festspeicher, von denen jeder ein $512 \cdot 8$ Bit-Matrix für total 4096 Bits enthält. Selbstverständlich kann ein einziger Festspeicher mit einer Kapazität von 8192 Bits an Stelle der zwei Festspeicher 12 und 14 verwendet werden.

Die Festspeicher 12 und 14 werden von einer Einschalt-dauer-Adressierschaltung 16 getaktet, welche die genaue Taktsequenz auf den Leitungen 18 und 19 den Festspeichern 12, 14 zugeleitet, um ein reihenförmig binär-gewichtetes Einschalt-dauerparameter-Steuersignal zu erzeugen. Die Adressierschaltung 16 ist an eine Taktgeberschaltung 20, die ein Rechtecksignal mit einer Frequenz von 20 kHz erzeugt, angeschlossen. Von den Rechtecksignalen werden durch die Adressierschaltung 16 Gruppen zu fünfzehn (15) Pulsen abgetrennt, die dann weiter in Zeitabschnitte von 8, 4, 2 und 1 Taktimpuls unterteilt werden. Für jede Gruppe von 15 Pulsen, welche die Adressierschaltung 16 empfängt, gibt sie während der acht und vier Zeitabschnitte ein hohes Ausgangssignal an die Leitung 18 und während der acht und zwei Zeitabschnitte ein hohes Ausgangssignal an die Leitung 19 ab.

Die durch die Festspeicher 12 und 14 erzeugten reihenförmigen, binär-gewichteten digitalen Steuerparameter umfassen mit Vorteil vier Auslösungsbits. Mit anderen Worten die Festspeicher 12 und 14 enthalten vier Informationsbits für jeden Phonemparameter und erzeugen dadurch 24 oder 16 mögliche Werte pro Parameter. Um den vier Bits ihre geeignete Binärstelle zu geben, werden das erste oder höchstwertige der vier

reihenförmigen Ausgangsbits im Steuerparameter, wenn beide Signale auf den Leitungen 18 und 19 hoch sind, das zweite Bit, wenn das Signal auf der Leitung 19 niedrig und das Signal auf der Leitung 18 hoch ist, das dritte Bit, wenn das Signal auf der Leitung 19 hoch und das Signal auf der Leitung 18 niedrig ist, und das vierte oder geringstwertige der vier Bits, wenn die Signale auf beiden Leitungen 18, 19 niedrig sind, erzeugt. Somit wird das erste oder höchstwertige Bit bei einer Periode von acht Taktimpulsen, das zweite Bit bei einer Periode von vier Taktimpulsen, das dritte Bit bei einer Periode von zwei Taktimpulsen und das vierte Bit bei einem Taktimpuls erzeugt. Auf diese Weise kann ein Analogsignal digital dargestellt werden.

Die Technik ist bekannt, um jedoch die Arbeitsweise des vorliegenden Synthetisators besser verständlich zu machen, werden die durch die Festspeicher 12 und 14 auf den Leitungen 22 erzeugten individuellen Steuerparameter kurz beschrieben.

Die Steuersignale F_1 und F_2 bestimmen die Stellungen der Resonanzpole in den ersten zwei Resonanzfiltern im Sprechtrakt 60. Die Steuersignale F_3 und F_4 bestimmen die Stellungen der Resonanzpole in dem dritten und vierten veränderlichen Resonanzfilter im Sprechtrakt 60. Der Nasallautverschluss und die Nasalfrequenzsteuersignale werden immer dann erzeugt, wenn stimmhafte Phoneme u, m oder ng vorliegen, um die im Sprechenergiezentrum auftretende Energie zu verringern, wenn diese stimmhaften Phoneme gesprochen werden. Insbesondere das den Nasallautverschluss steuernde Steuersignal bestimmt den zu entfernenden Sprechenergiewert und das Nasalfrequenzsteuersignal bestimmt die Frequenz bei der die Energie zu verringern ist. Die Reibelautfrequenz und Reibelauttieftpasssignale arbeiten miteinander und werden immer dann erzeugt, wenn Phoneme mit Reibelautenergie, wie f und s vorhanden sind. Diese Signale dienen zur spektralen Formung der Reibelaut-Erregungsenergie vor der Eingabe in den Sprechtrakt 60. Das Reibelauttieftpasssignal bestimmt insbesondere die Frequenz oberhalb welcher die Breitband-Reibelautanregungsenergie unterdrückt wird und das Reibelautfrequenzsignal bestimmt die Frequenz, bei welcher die maximale Stimmlosenergie auftreten wird. Das Übergangsgeschwindigkeitssteuersignal wird für jedes Phonem erzeugt und stellt zusammen mit dem Ausgangssignal aus der Geschwindigkeitssteuerschaltung die Übergangsgeschwindigkeit zwischen den eingeschwungenen Zuständen der vorstehend genannten Steuersignale ein. Das Taktsignal wird ebenfalls für jedes Phonem erzeugt und stellt zusammen mit dem Ausgangssignal aus der Geschwindigkeitssteuerschaltung 30 den Zeitraum zur Erzeugung für jedes Phonem ein. Ein Vokalamplitudensignal wird immer dann erzeugt, wenn ein Phonem mit einer stimmhaften Komponente vorhanden ist. Dieses Steuersignal steuert die Intensität der stimmhaften Komponente im Audioausgang. Das Vokalverzögerungssignal wird während bestimmten phonetischen Übergängen Reibelaut-Vokale erzeugt, wobei die Amplitude des Reibelautkomponenten rapid vergrössert wird. Das Vokalverzögerungssignal identifiziert jene Fälle, bei denen es erwünscht ist, dass der Sprechverzögerungsgenerator die Übertragung des Sprechamplitudensignals verzögert. Das Verschlusssteuersignal wird zur Simulation der Phonemwechselwirkung verwendet, die z. B. während der Erzeugung des Phonemes «b», welchem das Phonem e folgt, erzeugt wird. Insbesondere bewirkt das Verschlusssteuersignal, wenn es erzeugt wird eine abrupte Amplitudenmodulation im Audioausgang, welche den Aufbau und sofortige Freigabe der Energie, die während der Aussprache derartiger Phonemkombinationen auftritt, simuliert. Das Vokalspektralumrissignal ist ein anderes Steuersignal, welches das Sprechenergiespektrum spektral formt. Insbesondere steuert das Vokalspektralumrissignal einen ersten Tiefpassfilter, der die in den Sprechtrakt 60 eingebrachte Energie mit der maximalen Überlagerung, die beim Auftreten von reinen stimmlosen Phonemen auftritt, überlagert. Das Steuersi-

gnal F2Q verändert das Q oder die Bandbreite des zweiten Resonanzfilters F2 im Sprechtrakt und wird primär im Zusammenhang mit der Frequenz der Nasalphoneme n, m und ng verwendet. Nasalphoneme stellen einen höheren Energiebetrag beim ersten Formant (F1) und einen wesentlich kleineren und ausgedehnteren Energieinhalt bei höheren Formanten. Während des Vorhandenseins von Nasalphonemen wird somit das Steuersignal F2Q erzeugt, um das Q des Resonanzfilters F2 zu verringern, welcher wegen der kaskadenförmigen Anordnung der Resonanzfilter im Sprechtrakt 60, gültige Energiebeträge am Erreichen der höheren Formanten hindert. Das Reibelautamplitudensignal wird immer dann erzeugt, wenn ein Phonem mit einer stimmlosen Komponente vorhanden ist und wird zur Steuerung der Intensität der stimmlosen Komponente im Audiosignalausgang verwendet. Das Abschlussverzögerungssignal schliesslich wird während bestimmter phonetischer Übergänge selbst zu Reibelauten erzeugt, wobei es erwünscht ist, die Übertragung des Verschlusses, der Vokalspektralkontur des F2Q und der Reibelautamplitudensignale in der gleichen Weise, wie dies in Zusammenhang mit dem Sprechverzögerungssignal beschrieben würde, zu verzögern.

Die Ausgangssignalparameter aus dem Festspeicher 12 werden an eine erste Reihe von relativ langsam arbeitenden Übergangsfiltren 24 angelegt. Die Übergangsfiltren 24 sind zweckmässig aufgebaut und haben eine relativ lange Ansprechzeit im Verhältnis zu den eingeschwungenen Zuständen eines typischen Phonems, so dass abrupte Amplitudenänderungen in den Ausgangssignalen aus dem Festspeicher 12 unterdrückt werden. Die Übergangsfiltren 24 erzeugen graduale Wechsel zwischen den eingeschwungenen Pegeln der Signalparameter, um die glatten Übergänge zwischen den in der menschlichen Sprache vorhandenen Phonemen zu simulieren. Die Ansprechzeit der Übergangsfiltren 24 kann mittels des Ausgangssignals der Übergangssteuerschaltung 26 verstellt werden. Diese Steuerschaltung 26 kombiniert das Übergangsgeschwindigkeitssteuersignal aus dem Festspeicher 14 mit dem Ausgangssignal der Geschwindigkeitssteuerschaltung 30, um ein Rechtecksignal mit fester Frequenz und veränderbarer Impulsbreite zu erzeugen, dessen Einschaltdauer die Ansprechzeit der Übergangsfiltren 24 bestimmt.

Die die Geschwindigkeit wählenden zwei Bits aus dem Zwölf-Bit-Eingabebefehlswort werden direkt einer Phonemgeschwindigkeitssummierschaltung 28 zugeleitet. Diese Schaltung kombiniert die zwei Bits mit dem Sprechgeschwindigkeitssignal auf der Leitung 46 aus den Kennzeichenbefehl- und Steuerschaltung und gibt die Summe an die Geschwindigkeitssteuerschaltung 30 ab. Diese Steuerschaltung 30 erzeugt ein Rechtecksignal mit veränderbarer Pulsbreite, dessen Einschaltdauer zum Teil vom Wert des Ausgangssignals aus der Summierschaltung 28 abhängt. Das Signal aus der Steuerschaltung 30 wird der Übergangssteuerschaltung 26, den Phonemtakt-schaltung 32, einem Modulationsfilter 52 und einer zweiten Reihe von relativ langsam arbeitenden Übergangsfiltren 34 zugeleitet. Wie bereits erwähnt, kombiniert die Übergangssteuerschaltung 26 das Ausgangssignal aus der Geschwindigkeitssteuerschaltung mit dem Übergangsgeschwindigkeitssignal aus dem Festspeicher 14, um das Einschaltdauerübergangssignal, welches die Ansprechzeit der ersten Reihe von langsam arbeitenden Filtren 24 bestimmt, zu erzeugen.

Die Phonemtakt-schaltung 32, der auch das Sprechgeschwindigkeitsschalt-dauersignal der Steuerschaltung 32 empfängt, erzeugt ein Sägezahnsignal, das in einem die Dauer der Phonemerzeugung bestimmenden Zeitraum sich zwischen fünf und null Volt ändert. Der Anstieg des durch die Phonemtakt-schaltung 32 erzeugten Sägezahnsignals hängt sowohl von der Einschalt-dauer des Sprechgeschwindigkeitssignals aus der Steuerschaltung 30 als auch vom Wert des Phonemtakt-signals aus dem Festspeicher 14 ab. Es wird darauf hingewiesen, dass

das Phonemtakt-signal aus dem Festspeicher 14 den relativen Erzeugungszeitraum jedes einzelnen Phonems bewirkt, während die Steuerschaltung 30 die Totalgeschwindigkeit der Phonemerzeugung, d. h. die Sprechgeschwindigkeit, bestimmt.

Das Sprechamplitudensignal aus dem Festspeicher 14 wird an eine Modulationsschaltung 36 angelegt, welche das Steuersignal entsprechend dem Wert des Lautstärkensignals, das über die Leitung 48 aus der Schaltung 50 empfangen wird, moduliert. Das modulierte Steuersignal wird an einen Sprechverzögerungsgenerator 38 angelegt, der die Übertragung des Sprechamplitudensignals für eine bestimmte Zeit, die kürzer als die Dauer eines einzelnen Phonemzeitintervalls ist, verzögert, wenn immer ein Sprechverzögerungssignal durch den Festspeicher 14 erzeugt wird.

Das Reibelautamplitudensignal aus dem Festspeicher wird an eine Modulationsschaltung 40 angelegt, welche die Amplitude des Steuersignals entsprechend dem Wert des Lautstärkensignals, das über die Leitung 47 aus der Steuerschaltung 50 empfangen wird, moduliert. Das modulierte Steuersignal wird an einen Verschlussverzögerungsgenerator 42 angelegt, der gleich wie der Generator 38 arbeitet. Ausserdem werden auch der Verschluss, die Sprachspektralkontur und die F2Q-Steuersignale an den Verschlussverzögerungsgenerator 42 angelegt, der die Übertragung der vorstehend genannten Steuersignale für eine bestimmte Zeitperiode, die kürzer als die Dauer eines einzelnen Phonemintervalls ist, verzögert, wenn immer ein Verschlussverzögerungssignal durch den Festspeicher 14 erzeugt wird. Beachte, dass die durch die Generatoren 38 und 42 eingeführten Zeitverzögerungen von dem Zeitintervall des einzelnen Phonems, das erzeugt wurde, abhängen.

Wie vorstehend beschrieben stellen die Generatoren 38 und 42 sicher, dass die genaue Taktsequenz zwischen bestimmten Reibelaut-Vokal und Vokal-Reibelaut-Übergängen vorgesehen wird. Z. B. der Vokal-Reibelaut-Übergang beim Aussprechen des Buchstabens s enthält einen Vokalbestandteil, dessen Amplitude stark abfällt, während zur gleichen Zeit die Amplitude des Reibelautbestandteils normalerweise stark ansteigen würde. In diesem Fall wird der Generator 42 die Übertragung des Reibelautbestandteiles in Bezug auf den Vokalbestandteil verzögern, so dass der starke Anstieg im Reibelautenergiepegel nicht durch den starken Abfall des Vokalenergiepegels verloren geht. Die Erzeugung des Reibelautbestandteiles wird jedoch zeitlich etwas verzögert.

Die Ausgangssignale der Generatoren 38 und 42 werden an eine zweite Reihe relativ langsam arbeitender Übergangsfiltren 34, die die abrupten Amplitudenänderungen in den Steuersignalen, auf die gleiche Weise wie das vorher mit Bezug auf die Filtren 24 beschrieben wurde, glätten. Wie bei den ersten Filtren 24 wird die Ansprechzeit der zweiten Übergangsfiltren 34 auch durch das Sprechgeschwindigkeitseinschalt-signal aus der Steuerschaltung 30 gesteuert. Es wird jedoch darauf hingewiesen, dass die Ansprechzeit der zweiten Übergangsfiltren 34 nur durch das Sprechgeschwindigkeitssignal aus der Steuerschaltung 30 gesteuert wird, während die Ansprechzeit der ersten Übergangsfiltren 24 zusätzlich durch das Übergangsgeschwindigkeitssignal aus dem Festspeicher 12 gesteuert wird, und zwar weil die Übergangstaktung der Steuersignalparameter, die an die zweiten Übergangsfiltren 34 angelegt sind, nicht so kritisch wie die Taktung der Steuersignalparameter, die an die ersten Übergangsfiltren 24 angelegt sind, ist, und deshalb ist eine genaue Taktsteuerung durch das Übergangssteuersignal nicht erforderlich.

Die drei Bits für die Modulationswahl aus dem 12-Bit-Eingabebefehlswort werden direkt an einen Modulationsfilter 52 angelegt, der die binär-gewichteten Bits in ein einzelnes analoges Modulationssignal wandelt. Zusätzlich glättet der Modulationsfilter 52 die abrupten Amplitudenänderungen im Modulationssignal, und zwar auf die gleiche Weise wie es vorstehend

mit Bezug auf die Filter 24 und 34 beschrieben wurde.

Der Ausgang des Filters 52 ist an eine Sprechanregungsquelle 54 angelegt, welche die stimmhafte Erregungsenergie oder Stimmritzenwellenform erzeugt. Das Ausgangssignal des Filters 52 bestimmt die Höhe der Sprachenergie, welche der Grundfrequenz F_0 der Stimmritzenwellenform entspricht. Im bevorzugten Ausführungsbeispiel des Erfindungsgegenstandes weist die durch die Sprechanregungsquelle 54 erzeugte Stimmritzenwellenform im wesentlichen eine Sägezahnwellenform mit einem abgetrennten negativen Abschnitt des Signals auf. Diese neue Stimmritzenwellenform simuliert die Funktion der menschlichen Stimmritze besser, so dass die Natürlichkeit der erzeugten Sprache verbessert wird.

Um einen gewissen Grad von automatischer Modulationssteuerung, die bei bekannten Vorrichtungen bisher nicht verfügbar ist, zu erzielen, wird die durch die Sprechenergiequelle 54 erzeugte Grundfrequenz der Stimmritzenwellenform zusätzlich veränderbar ausgeführt, und zwar in Abhängigkeit von der F_1 , dem Nasalverschluss, der Sprechamplitude, dem Verschluss und den Reibelautamplitudensignalen. Insbesondere die vorstehend genannten Steuersignale werden einer Parametermodulationsschaltung 56 zugeleitet, die den Effekt der Steuersignale auf die Grundfrequenz der Stimmritzenwellenform modulieren.

Die Ausgangssignale der Filter 24 und 34 werden einer Reihe von Analog-Einschaltdauer-Wandlern 58 zugeleitet. Die Wandler 58 enthalten eine Mehrzahl Komparatoren, von denen jeweils ein Eingang ein 20-kHz-Dreieckssignal aus der Taktschaltung 20 und der andere Eingang eines der Steuersignale aus den Übergangsfiltren 24 und 34 empfängt. Die Komparatoren erzeugen Rechtecksignale fester Frequenz und veränderlicher Pulsbreite, deren Einschaltdauer dem Wert der an ihren Eingängen empfangenen Ergänzungssteuersignale sind.

Die Filter F_1 , F_2 , F_3 und F_4 , die in der Stimmritze erzeugte Überlagerung und Verschlusseinschaltdauersignale aus den Wandlern 58 werden direkt an den Sprechtraktfilter 60 angelegt. Der Filter 60 besteht im wesentlichen aus fünf hintereinandergeschalteten Resonanzfiltern, von denen vier variabel sind, einem Analog-Gatter und einem 20-kHz-Filter. Das Gatter spricht auf das Verschlusseinschaltsignal an, um die Amplitude des Audioausgangs zu modulieren und der 20-kHz-Filter unterbindet den Einfluss des Taktsignals auf den Audioausgang. Die variablen Resonanzfilter F_1 , F_2 , F_3 erzeugen die ersten drei Resonanzformanten im Energiespektrum des Audioausgangs und sind durch ihre entsprechenden Einschaltdauersignale abstimbar. Der Resonanzfilter F_1 ist über einen Frequenzbereich von 250 bis ca. 800 Hz abstimbar. Der Resonanzbereich F_2 ist über den Frequenzbereich von 760 bis 2400 Hz abstimbar. Der Resonanzfilter F_3 ist über den Frequenzbereich von 1200 bis 2550 Hz abstimbar. Wie nachfolgend beschrieben ist der Resonanzfilter F_4 , der den vierten Formant im Audioausgang erzeugt, ebenfalls variabel ausgeführt, ohne dass dazu ein zusätzliches Steuersignal erforderlich ist. Der Resonanzfilter F_4 wird durch das gleiche Steuersignal wie der Resonanzfilter F_3 abgestimmt und ist über einen Frequenzbereich von 2400 bis 3700 Hz veränderbar. Der Resonanzfilter F_5 ist ein Filter mit einem festeingestellten Resonanzpol, der einen fünften Formantpegel in den Audioausgang einführt, und zwar bei ca. 4400 Hz.

Die Filter F_2 und F_3 im Sprechtraktfilter 60 werden mit der stimmlosen Anregungssignalmenge aus einer Reibelautanregungssteuerung 62 gespeist. Nur die Filter F_2 und F_3 empfangen Reibelautenergie, weil festgestellt wurde, dass es ausreicht, dass die Reibelautenergie nur an diese zwei Punkte im Sprechtrakt gespeist werden, um das Frequenzspektrum aller Reibelautphoneme genau zu simulieren. Die Steuerung 62 empfängt das stimmlose oder Reibelautanregungssignal aus dem Reibelautgenerator 65, welcher die stimmlose Phonemmenge der

menschlichen Sprache erzeugt. Die Steuerung besteht im wesentlichen aus einer Gruppe von Analogsteuervorrichtungen, welche die Amplitude, die Frequenz und die Tiefpass-Signalmerkmale des Reibelautanregungssignals in Übereinstimmung mit den Einschaltssignalen, die aus den Wandlern 58 empfangen werden, ändert. Der Generator 64 besteht aus statischen Rauschquellen, welche den Klang von durch eine eingeschränkte Öffnung strömender Luft wie etwa beim Aussprechen der Phoneme s, f und n simuliert.

Die stimmhafte Signalmenge aus der Sprechanregungsquelle 54 wird auch dem Filter 60 zugeleitet, und zwar über eine Sprechanregungssteuerung 66. Die Steuerung besteht ebenfalls aus einer Gruppe von Analogsteuervorrichtungen, welche die Signalmerkmale des stimmhaften Anregungssignals in Übereinstimmung mit der Sprechamplitude, der Sprechspektralkontur, der Nasalfrequenz und den Nasalverschluss-einschaltdauersignalen, die von den Wandlern 58 empfangen werden, ändert.

Die in den Filter 60 eingespeiste Sprechenergie enthält einen zusätzlichen Formant, der durch eine Kompensationsschaltung 68 zum stimmhaften Anregungssignal addiert wird, um die Spektralenergie des Signals bei hohen Frequenzen zu steigern. Die stimmhafte Menge wird einer Stimmritzen-Überlagerungsschaltung 70 zugeleitet, welche die Resonanzüberlagerung in den Sprechtrakt einführt, um die Öffnung der Stimmritze beim Sprechen zu simulieren. Die Überlagerungsschaltung 70 erzeugt ein Einschaltdauersignal, das die Resonanz der Filter F_1 , F_2 und F_3 dämpft. Die Schaltung 70 erzielt die maximale Dämpfung während jenen Abschnitten der Stimmritzenwellenform, die der offenen Stimmritze entsprechen. Da die menschliche Stimmritze nur während der Erzeugung stimmhafter Phoneme aktiv ist, wird die Schaltung 70 so aufgebaut, dass sie ihr Überlagerungseinschaltdauersignal nur während der Erzeugung stimmhafter Phoneme an den Sprechtrakt 60 weiterleitet, wie dies durch den Empfang eines Sprechamplitudensignals auf der Leitung 72 dargestellt ist.

Schliesslich wird darauf hingewiesen, dass das vorliegende Ausführungsbeispiel der Erfindung eine Kennzeichenbefehlskoderteil und -steuerschaltung 50 enthält, um eine Gesamtgeschwindigkeit und Amplitudensteuerung des Audioausgangs zu erzeugen. Wie im Zusammenhang mit Fig. 6 ausführlich beschrieben werden wird, kann die Gesamtsprechgeschwindigkeit und/oder Lautstärke des Audioausgangs programmgemäss geändert werden, und zwar durch «Anrufen» der Kennzeichenbefehlsschaltung durch sein vorgewähltes Sieben-Bit-Phonem «name» und Eintragen der gewünschten Geschwindigkeits- und/oder Lautstärkenänderungen über die Geschwindigkeitsauswahlbits bzw. Modulationsauswahlbits. Die Schaltung 50 kann ausserdem ein ausgesprochenes Ruhephonem in die Sprachform einführen, um die menschliche Sprache realistischer zu simulieren.

Wie vorstehend im Zusammenhang mit der Beschreibung des Blockschemas von Fig. 1 erwähnt, werden drei der zwölf Bits im Eingabebefehlswort der Programmierung und Steuerung der Modulation oder Pitsch des Audioausgangs zugeordnet. Die drei Modulationsbits verbessern die Sprechqualität durch Erhöhung der Vielfalt von diskreten Modulationspegeln bei der Programmierung. Um dies zu erreichen, werden, wie Fig. 2 zeigt, jedes der drei Eingabedatenmodulationsbit 2⁶, 2⁷ und 2⁸ an einen gewichtenden Widerstand R_1 , R_2 bzw. R_3 angelegt. Die Widerstände R_1 , R_2 , R_3 sind parallel zueinander geschaltet und über den Verbindungspunkt 78 an den Modulationsfilter 80 angeschlossen. Die Widerstandswerte der Widerstände R_1 , R_2 , R_3 sind so ausgewählt, dass acht Modulationspegel erzeugt werden. Insbesondere der Widerstand R_1 , an dem das kleinste geltende Bit 2⁶ angelegt ist, hat einen Wert, der gleich dem vierfachen Wert des Widerstandes R_3 ist, an dem das höchste geltende Bit 2⁸ angelegt ist, und gleich dem zweifachen Wert des Widerstandes R_2 ist, an dem das mittlere Bit 2⁷

angelegt ist. Somit ist ersichtlich, dass der Beitrag des Modulationsbits 2^8 zur Amplitude des Signals am Verbindungspunkt 78 das Zweifache des Bits 2^7 und das Vierfache des Bits 2^6 beträgt. Der Modulationsfilter 80 umfasst einen langsam ansprechenden Filter, dessen Ansprechzeit durch das Sprechgeschwindigkeitseinschaltdauersignal aus der Geschwindigkeitsteuerschaltung gesteuert wird. Die relativ lange Aussprechzeit des Filters 80 glättet die abrupten Amplitudenänderungen im Signal, das am Summenpunkt 78 auftritt, wenn der Zustand der Eingabemodulationsbit sich ändert.

Wie in der Beschreibung der Fig. 1 erwähnt, wird der Ausgang des Modulationsfilters der Sprechanregungsquelle zugeleitet, die die stimmhafte Basisphonemmenge analog zur schwingenden Stimmritze im menschlichen Sprechtrakt erzeugt. Die Sprechanregungsquelle besteht im wesentlichen aus einem Integrator 82 mit einem Verstärker 86. Der Ausgang des Filters 80 ist über einen Koppelwiderstand an den negativen Eingang des Integrators 82 geführt. Der negative Eingang des Integrators liegt über einen Widerstand R_4 an eine Vorspannung von +12 Volt. Der positive Eingang des Integrators 82 liegt über einen Widerstand R_5 an Masse und der Ausgang des Integrators 82 ist über einen Kondensator C_1 an den negativen Eingang zurückgeführt. Da die Integration einer konstanten Spannung einen Spannungsanstieg ergibt, ist daraus zu ersehen, dass der Integrator 82 ein abfallendes Signal erzeugt, dessen Neigung proportional zur Spannung an seinem negativen Eingang ist. Beachte auch, dass der Ausgang des Integrators 82 über eine Rückführschaltung bestehend aus Widerständen R_6 , R_7 und R_8 , einer Diode D_1 und einem Verstärker 84 zum negativen Eingang zurückgeführt ist. Der Zweck der Rückführschaltung besteht darin, den Ausgang des Integrators auf seine ursprüngliche Spannung zurückzustellen und einen neuen Zyklus zu starten. Somit erzeugt der Integrator 82 ein Sägezahnsignal, wie es im beiliegenden Diagramm dargestellt ist, dessen Frequenz von der Amplitude des Signals aus dem Filter 80 abhängt.

Aus dem Stand der Technik hat sich ergeben, dass stimmhafte Signalmengen in Sägezahnwellenform eine natürlich klingendere Sprache als jene in Impulsform erzeugen. Dies beruht primär darauf, dass Sägezahnwellenformen eine breitere Amplitudenverteilung bei geraden und ungeraden Harmonischen erzeugt. Die Basissägezahn-Wellenform versagt jedoch bei Betrachtung der drei Grundfunktionen der menschlichen Stimmritze; (1) das Öffnen der Stimmritze, (2) das Schliessen der Stimmritze und (3) die geschlossene Stimmritze. Um die Funktionen der Stimmritze genauer zu simulieren, wird die Stimmritzenwellenform weiter modifiziert, und zwar indem man den Ausgang des Integrators 82 über einen Widerstand R_{14} an den negativen Eingang eines Verstärkers 86 anschliesst. Der positive Eingang des Verstärkers 86 ist an Masse angelegt. Der Ausgang des Verstärkers 86 ist über eine Diode D_3 und einen Shunt R_{15} an seinen negativen Eingang zurückgeführt. Die Diode D_3 wirkt als Rückführungskurzschluss für jene Signale, die ihre Durchbruchspannung übersteigen. Der Shunt R_{15} bildet eine lineare Rückführung für die an den Widerstand R_{14} liegenden Signale, die negativ bezüglich der Masse sind. Der Verstärker 86 invertiert also das Signal aus dem Integrator 82 und beschneidet die Sägezahnwellenform durch Subtrahieren der unteren Hälfte des Signals, wie dies in dem beiliegenden Diagramm aufgezeigt ist.

Der Wert der Diode D_3 ist so gewählt, dass etwas mehr als die Hälfte der Sägezahnwelle abgeschnitten wird, d. h. dass der Ausgang des Verstärkers mehr als 50% des Signals umfasst. Es wurde mit einer Schaltung zur Änderung des Prozentsatzes in Übereinstimmung mit der Erzeugung unterschiedliche Phoneme experimentiert, jedoch die wegen der grossen Anzahl von zusätzlichen Schaltungen sich steigernde Komplexität steht der Einführung dieser Schaltungen in das Ausführungsbei-

spiel des Erfindungsgegenstandes entgegen. Die feste Wellenform hat sich als sehr geeignet für die meisten Anwendungszwecke erwiesen. Praktisch ausgedrückt: es gibt einige Fälle, bei denen eine Änderung in der Trennhöhe des Signals einen deutlichen Unterschied in der Qualität des Audioausgangs ergeben.

Ausserdem wird das Ausgangssignal des Verstärkers 86 durch Simulierung der drei Grundfunktionen viel besser an die Funktion der menschlichen Stimmritzen angenähert. Der ansteigende Abschnitt der abgeschnittenen Wellen simuliert das Schliessen der Stimmritze und der flache Abschnitt simuliert die geschlossene Stimmritze. Die resultierende Stimmritzenwellenform gibt für den Fall, dass die menschliche Stimmritze kurz nach dem die maximale Anregung auftritt, um den Sprechtonen in Abhängigkeit davon das freie Schwingen zu ermöglichen. Weil die abgeschnittene Welle die maximale Anregung erzeugt, wenn das Signal an seiner positiven Spitze seine Richtung umkehrt, ist ersichtlich, dass die Wellenform das «Ruhe»-Merkmal der menschlichen Stimmritzenfunktion durch Vorsehen einer inaktiven Periode (entsprechend der geschlossenen Stimmritze) kurz nach dem Abschnitt des Signals, bei dem die maximale Anregung auftritt, simuliert.

Die neue Stimmritzenwellenform gewinnt zusätzlich dadurch an Bedeutung, dass sie auch in Kombination mit der Stimmritzenüberlagerungsschaltung angewendet wird, um eine Stimmritzenüberlagerung von Sprechresonanzen, ähnlich jener, welche in der menschlichen Stimme auftreten, zu erzeugen.

Es ist bekannt, dass die Frequenz mit der die menschliche Stimmritze schwingt, nicht konstant bleibt. Die Änderungen in der Grundfrequenz oder «Pitch» der menschlichen Stimme können willkürlich oder unwillkürlich verursacht werden. Willkürliche Wechsel im Pitch können als solche Verschiebungen und Formen beschrieben werden, welche eine Person einer Aussage zuweist, um die Wichtigkeit eines bestimmten Wortes anzugeben oder eine gewisse Emotion zu übermitteln. Unwillkürliche Wechsel werden andererseits durch den Stimmritzenunterdruck- und Muskelbewegungen bewirkt, welche gewöhnlich beim Aussprechen von Vokalen und Konsonanten auftreten. So z. B. der unwillkürliche Wechsel in der Grundfrequenz der Stimme beim Ansprechen der englischen Worte «beat» und «bat». Die Grundfrequenz fällt beim englischen Wort «bat» konstant ab und steigt beim Wort «beat», und zwar weil das Phonem «e» im englischen Wort «beat» einen grösseren Muskelanspannungszustand während der Aussprache erforderlich macht als das Phonem «ae» im englischen Wort «bat». Da es oft der Fall ist, dass der Anwender eines Synthetisators die Modulationsbefehlsbits nicht verwendet, ist die erzeugte Sprache sehr unnatürlich.

Wird ein Synthetisator ferner primär als Text-Audio-Wandler angewendet, so wird die optimale Anwendung der Modulationsbefehlsbits extrem schwierig. Wenn jedoch eine unwillkürliche Modulationszuweisung im Aufbau des Synthetisators enthalten ist, wird die Modulationsprogrammierung vereinfacht, weil man sich primär den willkürlichen Grundfrequenzwechseln widmen kann.

Im Ausführungsbeispiel der vorliegenden Erfindung ist die Modulationszuweisung enthalten, und zwar durch Ändern des Eingangssignals zum Integrator 82. Fig. 3a zeigt die Beziehung zwischen der Grundfrequenz und der Stellung des ersten Formanten während der Erzeugung von Vokalphonemen. Aus der Betrachtung des Diagramms wird klar, dass sich die Grundfrequenz umgekehrt bezüglich der Wechsel in der Stellung des ersten Formanten ändert, und zwar über das Spektrum der angegebenen Vokalphoneme. Dieses wird durch Anschliessen des Ausgangs des Übergangsfilters F_1 über einen Widerstand R_{11} an den positiven Eingang des Integrators vorgenommen. Somit ist ersichtlich, dass, wenn das Signal aus dem Filter F_1

ansteigt, sich die Differenz zwischen den Spannungen am positiven und negativen Eingang des Integrators 82 verringert. Somit wird die Neigung der abfallenden Flanke der Sägezahnwelle am Ausgang des Integrators 82 flacher, wodurch die Sägezahnwellenform gestreckt und damit die Frequenz des Signals verringert wird. Folglich wird die Grundfrequenz der Stimmritzenwellenform bezüglich der Änderungen im Signal aus dem Filter F1, welches die Stellung des ersten Formanten steuert, automatisch invers geändert.

Fig. 3b zeigt die Stellung der Hauptgrundfrequenz bei verschiedenen Konsonanten. Wie im Diagramm dargestellt, fällt die Hauptgrundfrequenz während des Vorhandenseins von Nasalphonemen, wie n, m oder «ng». Diese charakteristische Modulationsänderung wird bei dem vorliegenden Ausführungsbeispiel der Erfindung durch Anschliessen des Ausgangs aus dem Nasalübergangsfiler über einen Widerstand R10 an den positiven Eingang des Integrators 82 angewendet. Auf diese Weise verringert sich die Grundfrequenz der Stimmritzenwellenform, die am Ausgang des Integrators 82 erscheint, wenn ein Nasalsignal vorhanden ist. Insbesondere bewirkt die wegen des Vorhandenseins eines Signals aus den Nasalübergangsfiler erhöhte Spannung am positiven Eingang des Integrators 82 einen Abfall in der Neigung der abfallenden Flanke der Sägezahnwellenform. Somit fällt die Grundfrequenz der Stimmritzenwellenform, wenn Nasalphoneme erzeugt werden.

Wie Fig. 3b weiter zeigt, steigt die Grundfrequenz der Stimmritzenwellenform beim Auftreten von Reibelautphonemen wie «f», «h», «s» oder «sh» an. Um diese Modulationsmerkmale in den Synthetisator auszuführen, wird der Ausgang des Reibelautamplitudenübergangsfilters über einen Widerstand R9 an den negativen Eingang des Integrators 82 angeschlossen. Da ein Reibelautamplitudensignal immer auftritt, wenn ein Reibelautphonem erzeugt wird, steigt beim Auftreten eines Reibelautphonems die Spannung am negativen Eingang des Integrators 82. Durch das Ansteigen der relativen Spannung am negativen Eingang des Integrators 82 wird die Zeitkonstante der Schaltung verringert. Die Neigung der abfallenden Flanke der Sägezahnwellenform steigt dadurch und folgedessen steigt die Grundfrequenz des Ausgangssignals. Während der Erzeugung von Reibelautphonemen steigt somit die Grundfrequenz der Stimmritzenwellenform.

Schliesslich zeigt Fig. 3 noch, dass beim Auftreten von Phonem, wie «b», «d» oder «g», die Hauptgrundfrequenz fällt. Um diese Modulationsmerkmale auszuführen, ist es erforderlich die Steuersignalkombinationen auszuwählen, welche das Auftreten dieser Phoneme einheitlich angeben. Die Phoneme b, d und g sind Verschlussphoneme, welche die Erzeugung eines Verschlusssignals erforderlich machen. Ein Verschlussignal wird jedoch auch für die Verschlussphoneme p, t und k erzeugt. Wie das Diagramm in Fig. 3 zeigt, ist die Hauptgrundfrequenz für die Phoneme p, t und k wesentlich höher als die für die Phoneme b, d und g. Deshalb ist es erforderlich ein anderes Steuersignal zu erzeugen, um zwischen diesen zwei Gruppen von Phonemen unterscheiden zu können. Die Phoneme b, d und g sind stimmhafte Verschlussphoneme, während p, t und k stimmlose Verschlussphoneme sind. Durch Verwendung des Ausgangs des Sprechamplitudenübergangsfilters und seiner logischen Addierung zum Ausgang des Verschlussübertragungsfilters kann das Auftreten der Phoneme b, d und g ausgezeichnet bestimmt werden. Diese Modulationsmodifikation wird durch Anschliessen des Ausgangs des Sprechamplitudenübergangsfilters an den positiven Eingang des Integrators 82, wobei zwischen diesen zwei in Linie geschaltete Widerstände R12, R13 geschaltet sind, und am Ausgang des Verschlussübergangsfilters ein den Mittelpunkt der Widerstände R12, R13 ausführen, wobei eine Diode D2 zwischen diesen geschaltet ist. Hat der Widerstand R13 einen wesentlich höheren Widerstandswert als der Widerstand R12, so arbeitet die Schaltung als ein logi-

sches UND-Gatter. Dadurch steigt die Spannung am positiven Eingang des Integrators 82 nur an, wenn von beiden Filtern ein Ausgangssignal anliegt. Somit wird die Grundfrequenz der Stimmritzenwellenform verringert, wenn beide Ausgangssignale auftreten.

Es wird darauf hingewiesen, dass die Änderungen im Modulationsniveau des Audioausgangs, wie in der natürlichen menschlichen Sprache stufenweise auftreten, weil die Modulationsänderungsparameter von den Ausgängen der Übergangsfiler abhängen. Ferner da die vorstehend beschriebene automatische Modulationssteuerung weniger schwierig als programmierte Modulationsänderungen ist, wird das System primär als ein Text-Audio-Wandler angewendet, somit können die automatischen Modulationsänderungen durch einfache Änderung der Widerstandswerte der Schaltung offensichtlich gemacht werden.

Wie vorstehend bei der Beschreibung der Fig. 1 erwähnt, enthält der Synthetisator eine sogenannte «Hochpol»-Kompensationsschaltung, welche die Spektralenergie der Stimmritzenwellenform bei hohen Frequenzen erhöht. Der Ausdruck «Hochpol» bezieht sich auf die Formantenhöhe im Frequenzspektrum des hörbaren Bereichs. Jedoch ist bekannt, dass innerhalb des hörbaren Bereichs, die höherpoligen Formanten nichts zur Verständlichkeit des Audioausgangs beitragen. Es wurde jedoch festgestellt, dass ihr Vorhandensein die relative Spektralenergie beeinflusst, die bei niederen zur Sprachverständlichkeit beitragenden Formanten erreichbar ist. Demzufolge weist der Synthetisator eine Hochpol-Kompensationsschaltung auf, welche einen hohen Frequenzformanten der Stimmritzenwellenform bei ca. 4000 Hz hinzufügt. Dies wird dadurch erreicht, dass die abgeschnittene Stimmritzenwellenform, die am Ausgang des Verstärkers 86 erscheint, einem stark bedämpften Bi-Quad-Resonator 96 zugeleitet wird. Der Resonator 96 ist ein Filter mit einem festen Pol, der mit den Resonanzfiltern im wesentlichen identisch ist, die im Sprechtrakt verwendet werden und in Fig. 5 im Einzelnen dargestellt sind, ausser dass der Ausgang aus dem Resonator 96 den Bandpassausgang eher öffnet als der Ausgang des Tiefpass wie im Sprechtrakt. Der Ausgang aus dem Resonator 96 wird über einen Summierwiderstand R21 einem Summierpunkt 94 zugeleitet. Ebenso wird die abgeschnittene Stimmritzenwellenform über einen Summierwiderstand R20 dem Summierpunkt 94 zugeleitet. Das am Summierpunkt 94 auftretende Signal enthält somit die abgeschnittene Stimmritzenwellenform mit einem dazu addierten Formanten bei ca. 4000 Hz, welches die Spektralenergie der Wellenform bei hohen Frequenzen wirksam steigert.

Es wird darauf hingewiesen, dass die Stelle des Formanten, der zur Stimmritzenwellenform addiert ist, in der Frequenz tiefer liegt als der höchste Resonanzformant im Sprechtrakt. Es wurde festgestellt, dass diese Beziehung besonders wichtig ist, um die Qualität der erzeugten Sprache zu verbessern und bessere Resultate zu erhalten, als wenn der zur Stimmritzenwellenform addierte Formant der höchste Formant im Sprechsystem war.

In Fig. 4 ist die Wirkung des Resonators 96 auf die Spektralenergie der Stimmritzenwellenform graphisch dargestellt. Die Kurve I stellt die Spektralenergie der Stimmritzenwellenform ohne den Einfluss des Resonators und die Kurve II die Spektralenergie der Stimmritzenwellenform, die vom Resonator beeinflusst ist, dar. Aus diesem Diagramm kann ersehen werden, dass die Spektralenergie der Stimmritzenwellenform ohne Hochpol-Kompensation bei hohen Frequenzen stark abfällt. Mit dem Hinzufügen der Kompensation jedoch wird die Spektralenergie der Stimmritzenwellenform auf einen Wert oberhalb 4000 Hz, die der Resonanzfrequenz des Resonators entspricht, gehalten.

Es wird darauf hingewiesen, dass die Hochpol-Kompensa-

tion insbesondere für Synthetisatoren, wie sie in Fig. 1 dargestellt und in Zusammenhang mit dieser beschrieben sind, wichtig sind, bei denen der Sprechtrakt kaskaden- oder reihenmässig geschaltete Resonanzfilter aufweist. Dies wegen des Eigenverlustes des Anregungssignals, wenn es den niederfrequenten Resonator im Sprechtrakt durchläuft.

Zurück zu Fig. 2. Bevor das Sprechanregungssignal der Sprechanregungssteuerschaltung zugeleitet wird, wird es durch ein Hochpass-Filter 98 geführt, welcher die Frequenzen unterhalb ca. 150 Hz ausfiltert. Der Zweck des Hochpass-Filters 98 ist es, die Energie am niederfrequenten Ende des Spektrums aus der Stimmritzenwellenform zu eliminieren. Dies bewirkt die Eliminierung der Brummtöne aus dem Signal und führt zu schärferen hochfrequenten Abschnitten der Stimmritzenwellenform im Trakt. Dadurch wird die Verständlichkeit der Sprache verbessert.

Wie bereits im Zusammenhang mit Fig. 1 beschrieben, enthält der Synthetisator eine Stimmritzenüberlagerungsschaltung, welche die verringerte Dämpfung simuliert, die bei offener Stimmritze im menschlichen Sprechtrakt reflektiert wird. Der Zweck der Überlagerungsschaltung kann wie folgt besser erklärt werden. Der menschliche Sprechtrakt ist nur zeitweise an einem Ende offen, mundseitig, aber am anderen Ende geschlossen, stimmritzenseitig. Ist die Stimmritze offen, so wird die Dämpfung im Sprechtrakt verringert, woraus sich eine Bedämpfung der Formantresonanzen ergibt. Dies ist das Merkmal des menschlichen Sprechtraktes zu dessen Simulation die Überlagerungsschaltung eingeführt ist. Zurück zu Fig. 2. Das Sprechanregungssignal aus dem Verstärker 82 wird über ein Analog-Gatter 88 und einer Spannungsteilerschaltung, bestehend aus Widerständen R18 und R19, dem positiven Eingang eines Komparators 92 zugeleitet. Am negativen Eingang des Komparators 92 liegt das Dreieck-Taktsignal von 20 kHz an. Der Komparator 92 gibt ein Signal an seinem Ausgang ab, wenn die Amplitude des an seinem positiven Eingang angelegten Signals die Amplitude des an seinem negativen Eingang angelegten Signals übersteigt. Somit gibt der Komparator 92 ein 20-kHz-Ausgangssignal variabler Pulsbreite ab, dessen Einschaltdauer direkt proportional zum Potential der an seinem positiven Eingang angelegten Stimmritzenwellenform ist. Wenn die Frequenz des an den negativen Eingang des Komparators 92 angelegten Taktsignals in der Grössenordnung von 200 grösser als die Frequenz der an den positiven Eingang angelegten Stimmritzenwellenform ist, wird festgestellt, dass das Ausgangssignal des Komparators 92 als eine schnelle Folge von Spitzen auftritt, die nach und nach in der Amplitude höher ansteigen als die Stimmritzenwellenform. Nähert sich die Stimmritzenwellenform seiner maximalen Amplitude tritt das Ausgangssignal des Komparators 92 im wesentlichen als ein konstantes Signal auf, das mit einer schnellen Folge von nach und nach enger beisammenliegenden abfallende Spitzen überlagert ist. Ferner wird festgestellt, dass während des Teiles der Stimmritzenwellenform, der die geschlossene Stimmritze darstellt, d. h. während des ebenen Teils der Wellenform, kein Impuls am Ausgang des Komparators 92 auftritt.

Das Stimmritzenüberlagerungseinschaltssignal, das am Ausgang des Komparators 92 abgegeben wird, wird einem Sprechtraktfilter 60 zugeführt. In Fig. 4 sind die F3 und F3-Bi-quad-Resonanzfilter des Sprechtraktes im einzelnen dargestellt. Wie das Schaltbild zeigt, liegt das Stimmritzenüberlagerungseinschaltssignal an ein Analog-Gatter 100 an, das mit einem Widerstand R22 in Reihe geschaltet ist. Diese Reihenschaltung ist über den Bandpasswiderstand R23 des Resonators F3 geschaltet. Ist das Gatter 100 leitend, d. h. das Stimmritzenüberlagerungseinschaltssignal ist Null, so ist der Widerstand R22 unendlich gross und hat auf die Bandbreite des Filters F3 keinen Einfluss. Steigt die Einschaltdauer des Stimmritzenüberlagerungssignal an, wird das Gatter 100 leitend, dadurch wird der effek-

tive Widerstandswert des Widerstandes R22 verringert, welches die Bandbreite des Filters vermindert. Nähert sich das Stimmritzensignal seiner maximalen Einschaltdauer, erreicht der Widerstand R22 seinen Nennwert. Dadurch wird die Bandbreite des Filters F3 auf seinen minimalen Wert verringert. Damit wird eine Bedämpfung der Resonanz bei offener Stimmritze erzielt, wobei die maximale Bedämpfung auftritt, wenn die Stimmritzenwellenform seine positive Spitze erreicht, was der maximal offenen Stimmritze entspricht.

Es wird darauf hingewiesen, dass das Stimmritzenüberlagerungseinschaltssignal auch an zwei gleiche Gatter angelegt wird, die in der gleichen Weise wie beim Resonator F3 über die Bandpassabschnitte der Resonatoren F1 und F2 geschaltet sind. Somit wird die Bandbreite der drei Resonatoren F1, F2 und F3 durch die Stimmritzenwellenform verändert. Im Gegensatz zur Funktion der Sprechanregungsquelle des vorliegenden Synthetisators ist die Stimmritze im menschlichen Sprechsystem nicht über 100% der Zeit aktiv. Insbesondere während der Erzeugung stimmloser Phoneme ist die Stimmritze nicht vollständig aktiv. Deshalb ist es wünschenswert, die Stimmritzenüberlagerung nur während der Erzeugung stimmhafter Phoneme wirken zu lassen (Fig. 3). Dies wird durch Anlegen des Ausgangssignals der Sprechamplitudenübergangsfilter an den positiven Eingang eines Komparators 90 erreicht. Der negative Eingang des Komparators 90 ist an den Mittelpunkt einer Spannungsteilerschaltung bestehend aus den Widerständen R16, R17 angeschlossen. Die Spannungsteilerschaltung ist mit dem Widerstand R16 an eine +8 Volt Spannungsquelle angeschlossen und liegt mit dem Widerstand R17 an Masse. Der Ausgang des Komparators 90 ist an das Gatter 88, das als Schalter arbeitet, angeschlossen. Der Komparator 90 gibt ein Signal, welches das Gatter leitend macht, nur dann ab, wenn ein Signal aus dem Sprechamplitudenübergangsfilter vorliegt. Da das Vorliegen eines Sprechamplitudensignals das Vorhandensein eines stimmhaften Phonems anzeigt, ist das Gatter 88 leitend und es wird ein Stimmritzenüberlagerungseinschaltssignal nur während der Erzeugung stimmhafter Phoneme erzeugt. Demzufolge ist das Gatter 88 während der Erzeugung stimmloser Phoneme gesperrt und somit wird das Stimmritzenübergangseinschaltssignal unterdrückt.

Zurück zu Fig. 5. Es wird darauf hingewiesen, dass neben den variablen Filtern F1, F2 und F3 auch der Filter F4 variabel ausgeführt wurde, um die Natürlichkeit der erzeugten Sprache zu verbessern. Bekannte Synthetisatoren enthalten vier oder fünf Filter. Nur die ersten drei Filter F1, F2, F3 sind jedoch für gewöhnlich variabel. Es wurde jedoch festgestellt, dass eine Verstellbarkeit für den vierten Filter wünschenswert sein könne. Es wird gewöhnlich wegen der Komplexität bei der Erzeugung von zusätzlichen Daten oder Parametern nicht durchgeführt. Es wurde jedoch festgestellt, dass die Verschiebung des dritten und vierten Resonanzpoles bis zu einem bestimmten Grad gleich verläuft. Beim vorliegenden Synthetisator wird die Verstellung des vierten Resonanzfilters durch Einführen eines Gatters 106 und eines Widerstandes 103 zum Resonator F4 und Vorsehen eines Einschaltssignals für den Resonator F3, welches die Stelle des Frequenzpoles im Resonanzfilter F4 ohne das zusätzliche Daten oder Parameter erzeugt werden, müssen einfach durch Verwendung des variablen Steuersignals, welches den Resonanzfilter F3 verstellt, verstellbar gemacht werden kann.

In Fig. 6 ist ein Schaltschema der Kennzeichenbefehlskodier- und Steuerschaltung 50 dargestellt. In der Schaltung von Fig. 6 sind auch die Schaltungen zur Phonemhäufigkeitssummiierung und zur Modulation der Sprech- und Reibelautamplituden enthalten.

Wie in der Beschreibung von Fig. 1 erwähnt, weist jedes Phonem eine ihm zugeordnete vorausgewählte Zeitperiode auf, wie sie durch das Phonemtaktsignal, welches die Normal-

zeit, während welcher das Phonem erzeugt wird, angibt, bestimmt. Für den Fall, dass die Zeit eines gegebenen Phonems programmässig geändert werden soll, sieht der Synthetisator zwei Wertauswahlbits vor, welche dem Programmierer die Option von vier unterschiedlichen Zeiträumen für jedes gegebene Phonem geben. Werden die Bits nicht verwendet, wird jedes Phonem in seinem natürlichen Zeitraum erzeugt. Falls der Programmierer die Dauer eines gegebenen Phonems entweder verlängern oder verkürzen möchte, kann er dies über die zwei Bits vornehmen. Bei der Text-Audio-Wandlung kann das gleiche Phonem in der Programmsteuerung durch Änderung der Bits stärker oder schwächer angegeben werden.

Es ist eine programmierbare Wertsteuerung vorgesehen, die durch Anlegen der zwei Wertbits 2^9 und 2^{10} an einen Summierpunkt 128, und zwar über zwei Wertigkeitswiderstände R30 bzw. R31 durchgeführt wird. Der Widerstand R30 hat den zweifachen Wert des Widerstandes R31, somit wird das 2^9 -Bit das niederwertige Bit und das Bit 2^{10} das höherwertige Bit. Der Summierpunkt 128 liegt über einen Widerstand R28 an eine +8 Volt Spannungsquelle an. Der Ausgang des Summierpunktes 128 liegt am negativen Eingang eines Summiervverstärkers 130, ist an Masse angelegt und sein Ausgang ist über den Widerstand R32 an den negativen Eingang zurückgeführt. Der Summiervverstärker 130 wirkt als Strom/Spannungs-Wandler, indem er ein analoges Ausgangssignal, dessen Wert proportional zum Strom am Summierpunkt 128 ist, abgibt. Sind beide Häufigkeitsauswahlbits auf den logischen Zustand 0 gesetzt, so bewirkt die über den Widerstand R28 anliegende Vorspannung +8 Volt einen minimalen Strom am Summierpunkt 128, welcher dann die minimale Spannung am Ausgang des Summiervverstärkers 130 bestimmt. Bei einem bevorzugten Ausführungsbeispiel beträgt diese Spannung 2,3 Volt. Der Ausgang des Summiervverstärkers 130 ist an eine Seite eines Potentiometers R33 angeschlossen. Die andere Seite des Potentiometers R33 liegt an Masse. Der Schleifer des Potentiometers R33 ist an den positiven Eingang eines Komparators 132 und über einen Kondensator C2 an Masse angeschlossen. Am negativen Eingang des Komparators 132 liegt das 20-kHz-Dreieck-Taktsignal an. Der Komparator 132 erzeugt ein 20-kHz-Ausgangssignal variabler Pulsbreite, dessen Einschaltdauer durch den Wert des an seinem positiven Eingang anliegenden Signal bestimmt wird. Das Potential des am positiven Komparatoreingang angelegten Signals kann entweder durch manuelles Einstellen des Potentiometers R33 oder durch Verändern des Stromes am Summierpunkt, der die Spannung am Potentiometer R33 verändert, verändert werden. Es ist somit deutlich, dass die logischen Zustände der zwei Geschwindigkeitsauswahlbits die Stromhöhe am Summierpunkt 128 einstellen, welche dann die Einschaltdauer des am Komparatorausgang erzeugten Signals bewirken. Beim bevorzugten Ausführungsbeispiel sind die Geschwindigkeitsauswahlbits normal auf einen logischen Zustand «01» gesetzt, um zwei Pegel «verlängern» und einen einzelnen Pegel «verkürzen» in der Einschaltdauer des Sprechgeschwindigkeitssignals zuzulassen. Zurück zu Fig. 1. Das Sprechgeschwindigkeitsschaltdauersignal aus dem Komparator 132 wird der Phonemtaktschaltung zugeleitet, wo es mit dem Phonemtaktsignal kombiniert wird, um den Antrieb der durch den Phonemtakter erzeugten Taktrampe zu bestimmen. Somit ist daraus ersichtlich, dass die zwei Geschwindigkeitsauswahlbits Mittel zur programmartigen Änderung der Taktung einzelner Phoneme bilden.

Die Kennzeichenbefehlskodier- und -steuerschaltung sieht den vorliegenden Synthetisator mit der Möglichkeit, die Gesamtlautstärke und Sprechgeschwindigkeit des Audioausgangs programmartig zu ändern. Das bevorzugte Ausführungsbeispiel (Fig. 6) ist so ausgeführt, dass es die invertierte Logik des Eingabebefehlswortes annimmt. Darunter ist zu verstehen, dass das Vorhandensein eines Signals durch einen logischen

Zustand «0» und das Fehlen eines Signals durch einen logischen Zustand «1» angezeigt wird.

Wie das Schaltbild in Fig. 6 zeigt, empfängt die Kennzeichensteuerschaltung alle zwölf Bits aus dem Datenbefehlswort. Der Schaltung ist eine einheitliche Sieben-Bit-«Nennung» zugeleitet, welche auf den sieben Eingangsleitungen eingegeben werden, um die Steuerschaltung anzurufen. Die sieben Phonemauswahlbits werden einer Logikschaltung 110 zugeleitet, welche die über die sieben Eingangsleitungen übertragene Information kodiert, wenn die Steuerschaltung angerufen ist. Im aufgezeigten bevorzugten Ausführungsbeispiel ist der Steuerschaltung die Bezeichnung «0000000» zugeteilt. Somit arbeitet die Logikschaltung 110 als ein ODER-Gatter mit sieben Eingängen, welches an seinem Ausgang ein niedriges Ausgangssignal abgibt, wenn an allen Eingängen ein logischer Zustand «0» auftritt.

Ist die Steuerschaltung angerufen, so werden zwei der Modulationsauswahlbits 2^6 und 2^7 als Amplitudenauswahlbit angewendet und die zwei Geschwindigkeitsauswahlbits werden zur Änderung der gesamten Sprechgeschwindigkeit des Audioausgangs verwendet. Ausserdem wird das dritte Modulationsauswahlbit 2^8 als Stimmlosphonembit, dessen Funktion nachfolgend ausführlich beschrieben wird, verwendet. Jedes der vier Amplituden- und Geschwindigkeitsbits liegt an einem der vier NOR-Gatter 112, 114, 116, 118 an. Gibt die Logikschaltung 110 ein niedriges Ausgangssignal an die Leitung 111 ab, so werden die Gatter 112, 114, 116, 118 leitend. Mit anderen Worten ist die Kennzeichensteuerschaltung angerufen, so erscheint an den Ausgängen der Gatter 112, 114, 116, 118 der invertierte Zustand der Amplituden- und Auswahlbits. Die Ausgänge der vier NOR-Gatter liegen jeweils am D-Eingang eines J-K-Flip-Flops 120, 122, 124, 126 an.

Die Flip-Flops 120, 122, 124, 126 werden durch das Kennzeichentaktsignal, das über die Leitung 135 aus einem Exklusiv-NOR-Gatter 137 empfangen wird, getaktet. Ein Eingang des NOR-Gatters 137 liegt an Masse, der andere Eingang ist an den Ausgang des NOR-Gatters 134 angeschlossen. Ein Eingang des NOR-Gatters 134 ist an den Ausgang einer Verzögerungsschaltung 136 bestehend aus Widerstand R35 und Kondensator C4 und der andere Eingang ist an den Q-Ausgang eines J-K-Flip-Flops 154 angeschlossen. Die Verzögerungsschaltung empfängt ein verzögertes Kennzeichensignal, das über eine Leitung 131 von einer anderen Verzögerungsschaltung 138 bestehend aus Widerstand R34 und Kondensator C3 abgegeben wird. Diese Verzögerungsschaltung empfängt über die Leitung 111 das Ist-Kennzeichensignal aus der Logikschaltung 110.

Nimmt man nun an, dass der Q-Ausgang des Flip-Flops 154 auf einen logischen Zustand «0» gesetzt ist, so wird ersichtlich, dass wenn das Signal auf der Leitung 131 «0» wird, der Ausgang des NOR-Gatters 134 «1» wird und den Ausgang des NOR-Gatters 137 ebenfalls «1» werden lässt. Dieser taktet die Flip-Flops 120, 122, 124, 126 und trägt die am D-Eingang der Flip-Flops anliegenden Daten ein. Es wird darauf hingewiesen, dass die Verzögerungsschaltung 138 sicher stellt, dass die Daten der Amplituden- und Geschwindigkeitsauswahlbits vor dem Kennzeichentaktsignal auf der Leitung 135 an den D-Eingängen der Flip-Flops 120, 122, 124, 126 anliegen. Die Daten (invertiert) aus den Amplituden und Geschwindigkeitsbits werden somit an den Q-Ausgängen der vier J-K-Flip-Flops 120, 122, 124, 126 übertragen, wenn die Kennzeichenschaltung angerufen ist.

Die zwei die Sprechgeschwindigkeit steuernden Flip-Flops 124, 126 sind mit ihren Q-Ausgängen zwei Widerstände R27 bzw. R29 an den Summierpunkt 128 angeschlossen. Folglich wirkt auch der logische Zustand des Q-Ausgangs der Flip-Flops 124, 126 auf den Strom am Summierpunkt 128 ein, der, wie vorstehend beschrieben, die Spannung am Potentiometer R33 bestimmt.

Der Widerstandswert des Widerstandes R27 ist vorzugs-

weise halb so gross wie der des Widerstandes R29. Folglich hat der logische Zustand des Flip-Flops 124 eine grössere Wirkung auf den Strom am Summierpunkt 128 als der logische Zustand des Flip-Flops 126. Ausserdem liegt das Nullkodiersignal auf der Leitung 139 am Rückstelleingang R des Flip-Flops 124 und am Setzeingang S des Flip-Flops 126 an. Das Signal auf der Leitung stellt die Flip-Flops zurück. Demzufolge hat der Q-Ausgang des Flip-Flops 124 den Zustand «0» und der Q-Ausgang des Flip-Flops 126 den Zustand «1», welche als Grundstellung der Flip-Flops definiert werden. Somit sind durch diese Grundstellung zwei Pegel «erhöhen» und ein einzelner Pegel «vermindern» in der Gesamtsprechgeschwindigkeit des Audioausgangssignals erhältlich.

Beim bevorzugten Ausführungsbeispiel der vorliegenden Erfindung ist der Wert des Widerstandes R29 (der grösser als der des Widerstandes R27 ist) kleiner als der Wert des Widerstandes R31 (der den halben Wert des Widerstandes R30 hat). Auf diese Weise werden die Gesamtänderungen in der Sprechgeschwindigkeit des Audioausganges, die durch die Änderungen der logischen Zustände der Flip-Flops 124 und 126 bewirkt werden, wenn die Kennzeichenschaltung angerufen ist, besser ausgesprochen als die Schwankungen in den relativen Zeiträumen der einzelnen Phoneme, die durch Wechsel in den logischen Zuständen der Geschwindigkeitsauswahlbits 2⁹ und 2¹⁰, wenn sie über die Widerstände R30 bzw. R31 anliegen.

Da die logischen Zustände der Flip-Flops 124 und 126 erst bei Empfang eines Taktimpulses am Takteingang geändert werden, wird erkennbar, dass ein programmierter Wechsel in der Gesamtsprechgeschwindigkeit nach dem Kennzeichenphonem bestehen bleibt. Insbesondere sind die Änderungen in den logischen Zuständen der Flip-Flops 124 und 126 fest und erlauben die nachfolgenden Einstellungen in den zwei Geschwindigkeitsauswahlbits bis der Kennzeichenphonembefehl wieder eingetroffen ist. Somit schafft das vorliegende System die Möglichkeit einer programmierbaren Einstellung der Taktung der einzelnen Phoneme, und zwar über die Geschwindigkeitsauswahlbits oder einer programmierbaren Änderung der Gesamtsprechgeschwindigkeit des Audioausgangs, und zwar über die Kennzeichenbefehls- und -steuerschaltung in Kombination mit den zwei Geschwindigkeitsauswahlbits.

Die Q-Ausgänge der Lautstärkensteuer-Flip-Flops 120 und 122 im Amplitudensteuerteil in Fig. 6 sind über zwei Wertigkeitswiderstände R24 bzw. R25 an einen Summierpunkt 140 angeschlossen. Der Summierpunkt 140 liegt über einen Widerstand R26 auch an eine Vorspannung von +5 Volt. Demzufolge steuern die logischen Zustände der Flip-Flops 120 und 122 zusammen mit dem Zuwachs von +5 Volt Vorspannung den Strom am Summierpunkt 140. Die +5 Volt Vorspannung, wenn sie über den Widerstand R26 angelegt ist, stellt den minimalen Strom am Summierpunkt 140 ein, wenn an den Q-Ausgängen beider Flip-Flops ein logischer Zustand «0» auftritt. Der Summierpunkt 140 ist mit zwei Analoggattern 142 und 144 verbunden. An den Schalteingängen der Analoggatter 142 und 144 liegen die Reibelaut- und Vokalamplitudensignale aus dem Festspeicher 14 an. Die andere Seite der Gatter 142 und 144 ist jeweils an den positiven Eingang eines Operationsverstärkers 150 bzw. 152 angeschlossen, und zwar über zwei andere Gatter 146 bzw. 148, deren Funktion später beschrieben wird.

Die OP 150 und 152 sind jeweils in eine Spannungsführerschaltung geschaltet, und zwar mit ihren positiven Eingängen über Widerstände R36 bzw. R37 an Masse. Die Ausgänge sind an die negativen Eingänge zurückgeführt. Auf diese Weise bilden die Operationsverstärker 150 und 152 einen Treiberkreis mit geringer Dämpfung für die Verschluss- und Sprechverzögerungsschaltungen, an welchen die Ausgangssignale der Operationsverstärker 150 bzw. 152 angelegt sind.

Die Höhe der Reibelaut- und Vokalamplitudensignale wird

in der folgenden Weise moduliert. Der Summierpunkt 140 liefert eine konstante Spannung an die Gatter 142 und 144, dessen Höhe, z. B. 3 Volt, durch die logischen Zustände der Flip-Flops 120 und 122 bestimmt wird. Werden die Gatter 142 und 144 die Reibelaut- und Vokalamplitudensignale dauernd leitend gemacht, so bilden sie einfache Leiter. In diesem Fall wurde die konstante Spannung vom Summierpunkt 140 einfach unverändert an die positiven Eingänge der Operationsverstärker 150 und 152 weitergeleitet. Da jedoch die Reibelaut- und Vokalamplitudensignale aus dem Festspeicher eine Rechteckimpulswelle mit zeitgewichteter variabler Pulsbreite, das im Anschlag zwischen Null und 5 Volt variiert, umfasst, werden die Gatter 142 und 144 das Signal vom Summierpunkt 140 durchlassen, wenn das Amplitudensignal aus dem Festspeicher 14 hoch ist. Sind dagegen die Amplitudensignale aus dem Festspeicher 14 hoch, sind die Gatter 142 und 144 nichtleitend.

Wie Fig. 7 zeigt, sind die Ausgangssignale aus den Gattern 142 und 144 Rechtecksignale mit variabler Impulsbreite, deren Einschaltdauer zur Einschaltdauer der Reibelaut- und Vokalamplitudensignale äquivalent ist, deren Spannungsanschlag aber auf die Höhe der Spannung am Summierpunkt 140 begrenzt ist. Mit anderen Worten, die von den Gattern 142 und 144 abgegebenen Ausgangssignale sind den Reibelautamplituden- bzw. Vokalamplitudensignalen äquivalent, ausser dass die Amplitude des Signals moduliert wurde, um der Spannung am Summierpunkt 140 zu entsprechen. Da der Wert der Amplitudensignale durch ihren Amplitudenwert über 15 Taktimpulse bestimmt wird, wird somit die Lautstärke des Audioausgangs dementsprechend geändert. Da ausserdem die logischen Zustände der Flip-Flops 120 und 122 konstant bleiben, bis ein anderes Kennzeichentakt signal aus der Leitung empfangen wird, wird die Gesamtänderung in der Lautstärke des Audioausgangs aufrechterhalten bis ein nachfolgender Kennzeichenbefehl auftritt, unabhängig von den Änderungen in den zwei Modulationsauswahlbits 2⁶ und 2⁷.

Die Nullkodierleitung 139 ist an den Rückstelleingang des Flip-Flops 120 und den Setzeingang des Flip-Flops 122 angeschlossen. Da das Flip-Flop 120 das weniger wertige Bit und das Flip-Flop 122 das höher wertige Bit liefert, ist eine Normalamplitudeneinstellung mit zwei Pegeln «Erhöhung» und einem einzelnen Pegel «Verringern» möglich.

Wie vorstehend erwähnt, versieht die Kennzeichenbefehl- und -steuerschaltung das System mit der Fähigkeit, ein ausgesprochenes Ruhephonem in das Sprachmuster einzuführen.

Wie Fig. 6 zeigt, wird das Ausgangssignal aus der Logikschaltung 110 zusätzlich an einen der Eingänge einem NOR-Gatter 156 angelegt. Das Ruhephonembit 2⁸ liegt am anderen Eingang des NOR-Gatters 156 an. Der Ausgang des NOR-Gatters 156 ist an den D-Eingang des JK-Flip-Flops 154 angeschlossen. Demzufolge wird, wenn die Kennzeichensteuerschaltung angerufen ist, wie dies durch ein niedriges Ausgangssignal aus der Logikschaltung 110 angezeigt wird, und das Ruhephonembit 2⁸ einen logischen Zustand «0» hat, der Ausgang vom NOR-Gatter 156 hoch sein. Der Q-Ausgang des Flip-Flops 154 ist an den Schalteingang der Gatter 146 und 148 angeschlossen. Mit der Nullkodierleitung 139, die an den Rückstelleingang des Flip-Flops 154 angeschlossen ist, ist der Q-Ausgang des Flip-Flops 154 hoch. Somit sind beim Fehlen eines Ruhephonems die Reibelaut- und Vokalamplitudensignale durch die Gatter 146 bzw. 148 geleitet. Ist jedoch ein Ruhephonem präsent, wird ein hohes Signal an den D-Eingang des Flip-Flops 154 abgegeben, welches den logischen Zustand des Q-Ausganges «0» werden lässt, wenn das entsprechende Ruhephonemtakt signal auf der Leitung 158 an den Takteingang des Flip-Flops 154 anliegt. Tritt dieses ein, so bleiben die Gatter 146 und 148 nichtleitend und verhindern die Übertragung sowohl der Reibelautamplitudensignale als auch der Vokalamplitudensignale. Beim Fehlen beider Amplitudensignale werden weder stimmhafte noch

stimmlose Anregungssignale in den Sprechtrakt eingebracht.

Wie nachfolgend im Einzelnen beschrieben ist, obwohl die Dauer des Kennzeichenbefehlsphonems extrem kurz ist, ist die Dauer des Ruhephonems der Dauer eines gewöhnlichen stimmhaften Phonems äquivalent. Daraus folgt, dass das Aussprachemuster für jedes Phonem während einer Ruhephonemperiode, die dem Kennzeichenbefehl folgt, erzeugt werden kann. Die primären Vorteile dieses neuen Merkmals sind wie folgt.

Obwohl theoretisch viele Sprachklänge durch eine eigentliche Phonemkombination erzeugt werden sollten, sind tatsächlich gewisse Sprachklänge durch Anwendung von Phonemen nicht einfach herzustellen. Z. B. Worte, die mit einem Vokal beginnen, sowie Worte, die mit den Buchstaben «l» oder «w» beginnen, bei welchen die Aussprachemuster vorbereitet werden, bevor das tatsächliche Aussprechen des Wortes beginnt. Insbesondere beachte man, wie der Mund die Aussprache der englischen Worte «oak», «ear», «like» und «walk» vorbereitet, bevor das Wort tatsächlich gesprochen wird. Ohne diese Vorbereitung beginnen diese Worte zu abrupt und klingen unnatürlich, als wenn das erste Phonem in jedem Wort teilweise weggelassen wird.

Das Merkmal des Ruhephonems im vorliegenden Ausführungsbeispiel kann verwendet werden, um dieses Aussprachemerkmal der menschlichen Sprache durch Vorsehen der Einrichtung zum Einstellen des Aussprechmusters für ein bestimmtes Phonem bevor dieses Phonem tatsächlich erzeugt wird, zu simulieren. Wenn z. B. ein englisches Wort mit dem Buchstaben «w» beginnend zu erzeugen ist, würde die bevorzugte Eingabebefehlswortfolge ein Ruhekennzeichenphonem abrufen, welches von zwei «w»-Phonemen gefolgt wird. Obwohl das erste dem Kennzeichenbefehl folgende «w»-Phonem nicht ausgesprochen wird, wird auf diese Weise das Aussprechmuster für das «w»-Phonem während der Ruhephonemdauer doch geformt. Dadurch klingt das «w»-Phonem merklich weicher und natürlicher.

Ausserdem kann dieses Merkmal des Ruhephonems auch zur Verbesserung der Spracherkennung von bestimmten Klängen, die an den Enden der Worte auftreten, verwendet werden. Insbesondere Worte, deren Enden auslaufen wie jene, die mit Nasalphonemen enden, klingen als wenn ein zusätzliches Phonem eingeführt wurde, wenn das Aussprachemuster des letzten Phonems abrupt beendet wird. Z. B. wenn das «n»-Phonem im englischen Wort «sun» abrupt beendet wird, klingt dieses Wort mehr nach «suna». Dies beruht primär auf der Tatsache, dass die Restenergie im Sprechtrakt stimmhaft wird, und zwar etwas anders als ein «n» nach der Dauer des «n»-Phonems.

Um diese Erscheinung zu verhindern, kann der Ruhekennzeichenbefehl in Kombination mit einem zusätzlichen «n»-Phonem verwendet werden, um dem Ende des Wortes ein «Ruhe-n» hinzuzufügen. Auf diese Weise wird die Aussprachform des «n-Phonems» erhalten und bewirkt, dass das nasale «n» natürlicher klingt.

Wie vorstehend angesprochen, ist die relative Taktung der verschiedenen Takt- und Datensignale in der Kennzeichenschaltung für ihre richtige Funktion wichtig und wird daher im Detail beschrieben. Fig. 8 zeigt ein Signaldiagramm, welches die Zustände der verschiedenen Signale an ausgewählten Punkten in der dargestellten Kennzeichenschaltung darstellt. Wird das Kennzeichenphonem aufgerufen, um die Gesamtgeschwindigkeit und/oder Lautstärke des Audioausgangs zu verändern, ist es wünschenswert, sofort mit dem Phonem weiterzufahren, ohne auf einen ganzen Phonemzeitraum zum Kennzeichenbefehl zu vertrauen. Dies, weil die Kennzeichenschaltung den relativ langen Zeitraum, der der Produktion eines typischen Phonems zugeteilt ist, um die befohlenen Änderungen durchzuführen, nicht benötigt. Um eine Pause im Aussprachemuster zu verhindern, wenn immer die Gesamtge-

schwindigkeit und/oder Lautstärke des Audioausgangs geändert wird, wird somit die Kennzeichenschaltung ein anderes Phonemtaktsignal erzeugen, das dem Taktsignal, welches das Kennzeichenphonem abrufen, sofort folgt.

Das Phonemtaktsignal, welches die Taktung der Eingabebefehlsörter (PCI) steuert, wird zwei Exklusiv-ODER-Gatter 164 und 166 zugeleitet. Der andere Eingang des ODER-Gatters 166 ist an den Ausgang der Verzögerungsschaltung 138 angeschlossen, welche das verzögerte Kennzeichensignal (FD) an die Leitung 131 abgibt. Der andere Eingang des ODER-Gatters 164 ist auch an den Ausgang der Verzögerungsschaltung 318 angeschlossen, und zwar über einen Inverter 168 und eine zweite Verzögerungsschaltung 160. Das am Ausgang der zweiten Verzögerungsschaltung 160 auftretende Signal wird durch die Bezeichnung (FDD) gekennzeichnet.

Die Ausgangssignale aus den ODER-Gattern 164 und 166 werden einem weiteren Exklusiv-ODER-Gatter 170 zugeleitet, dessen Ausgang an einen Eingang eines NOR-Gatters 172 angeschlossen ist. Der andere Eingang dieses NOR-Gatters 172 ist an den Ausgang der Verzögerungsschaltung 138 angeschlossen. Der Ausgang des Gatters 192 liegt an einem anderen NOR-Gatter 174 an, das mit seinem anderen Eingang über einen Inverter 176 an die PCI-Leitung angeschlossen ist. Das Ausgangssignal 174 kann dem Phonemtaktsignal als gleich betrachtet werden.

Angenommen dass der Kennzeichenphonembefehl nicht auftritt, wie es durch hohes Ausgangssignal F aus der Logikschaltung 110 angezeigt wird, so wird das Phonemtaktsignal PCO durch die Taktschaltung nicht gewechselt. Mit anderen Worten das Phonemtaktsignal PCO ist gleich dem Phonemtaktsignal PCI. Unter diesen Bedingungen wird eine normale Taktung des Eingabebefehlswortes durchgeführt.

Beim Auftreten eines Kennzeichenphonems jedoch wird das Ausgangssignal der Logikschaltung 110 niedrig sein. Tritt dieses auf, so addiert die Taktschaltung einen zweiten Phonemtaktimpuls zu dem Phonemtaktsignal.

Nach dem Diagramm von Fig. 8 wird dies wie folgt durchgeführt. Da wir hier die invertierte Logik anwenden, erscheint der Phonemtaktimpuls auf der PCI-Leitung als ein negativer Puls von ca. 180 μ s Dauer. Die ansteigende Flanke des Taktsignals, wie zur Zeit t1 im Zeitdiagramm angezeigt, entspricht dem Zeitpunkt, wenn der Kennzeichenphonembefehl anfangs abgerufen wird. Nach einer Zeitverzögerung von ca. 110 μ s, zur Zeit t2, spricht die Logikschaltung 110 auf den Phonembefehl an und gibt ein niedriges Ausgangssignal an seinen Ausgang F ab. Diese 110 μ s Verzögerung tritt primär wegen der Eigenverzögerung im Tastenfeld auf, welches die digitalen Eingabebefehlsörter liefert. Ca. 10 μ s nach der Zeit t2, die durch die Verzögerungsschaltung 138 eingeführten Verzögerung gleicht, wird das verzögerte Kennzeichensignal FD am Ausgang der Schaltung 138 abfallen und bewirken, dass das Phonemtaktsignal PCO auch abfällt, wie das zur Zeit t3 aufgezeigt ist. Nach einem weiteren Zeitabschnitt von ca. 220 μ s, der gleich der durch die Verzögerungsschaltung 160 angeführten Verzögerung ist, wird das zweifach verzögerte und invertierte Kennzeichensignal FDD am Ausgang der Schaltung 160 abfallen und bewirken, dass das Phonemtaktsignal PCO wieder ansteigt. Somit wird zu der Zeit t6 eine andere ansteigende Flanke zum Phonemtaktsignal addiert, die das nächste Phonembefehlswort ca. 340 μ s nach dem der Kennzeichenphonembefehl abgerufen ist, abrufen.

Um jedoch sicherzustellen, dass die programmierten Änderungen in der Gesamtgeschwindigkeit und/oder Lautstärke des Audioausgangs ausgeführt werden, ist es wichtig, dass das Kennzeichentaktsignal FCL vor der Zeit t6 an die Leitung 135 abgegeben wird. Mit anderen Worten, die vier Geschwindigkeits- und Lautstärken-Flip-Flops 120, 122, 124 und 126 müssen während der 220 μ s dauernden Verzögerungszeit, die durch die

Verzögerungsschaltung 160 eingeführt wird, getaktet werden. Zur Zeit t_3 wird das verzögerte Kennzeichensignal FD auf der Leitung 131 einer weiteren Verzögerungsschaltung 136 zugeleitet, welches das Kennzeichensignal um weitere ca. 50 μs verzögert. Wenn dieses zweifach verzögerte Kennzeichensignal FD' zur Zeit t_5 abfällt (angenommen es fehlt ein Ruhephonem) steigt das Kennzeichentaktsignal FCL auf der Leitung 135 an, welches die am D-Eingang der Flip-Flops 120, 122, 124 und 126 ausstehende Information einträgt.

Ist die Kennzeichenschaltung angerufen, um ein Ruhephonem in den Sprechausgang einzuführen, so ist es erwünscht, dass die Dauer des Ruhephonems mit der Dauer eines typischen Phonem zusammenfällt. Ausserdem ist im bevorzugten Ausführungsbeispiel der Erfindung die Kennzeichenschaltung so angepasst, dass der Status der Geschwindigkeits- und Amplituden-Flip-Flops 120, 122, 124 und 126 aufrechterhalten wird, wenn ein Ruhephonem erzeugt wird, so dass die Bedingungen, die vor der Erzeugung des Ruhephonems bestehen, nach der Erzeugung des Ruhephonems weiter bestehen.

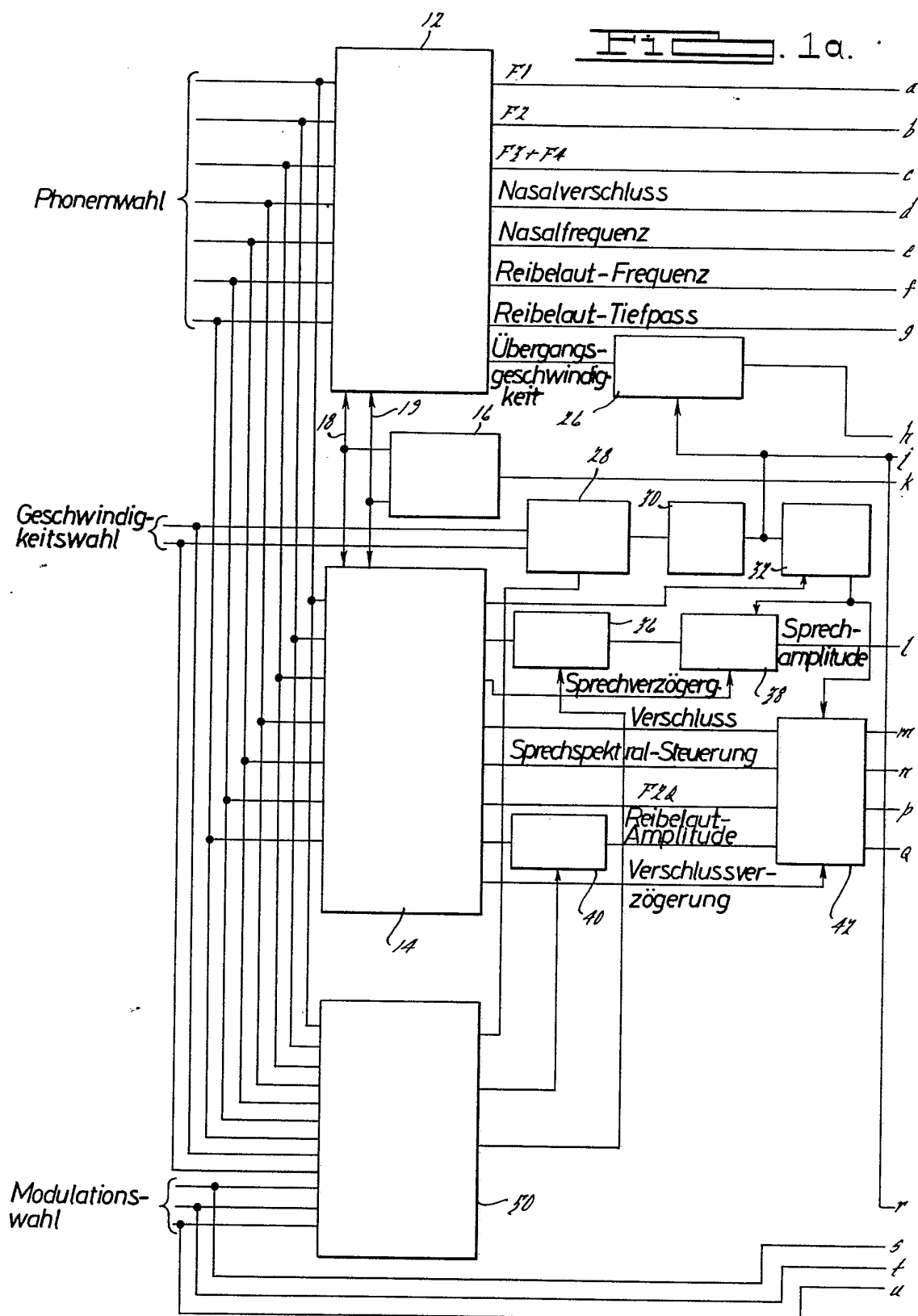
Zurück zu Fig. 6. Der Ausgang des Gatters 174 wird einer anderen Verzögerungsschaltung 162, bestehend aus dem Widerstand R 39 und dem Kondensator C6 zugeleitet. Der Ausgang der Schaltung 126 liegt an beiden Eingängen eines NOR-Gatters 176 an. Der Ausgang des Gatters 176 ist an den Takteingang CL des Ruhephonem-Flip-Flops 154 angeschlossen. Wie das Zeitdiagramm von Fig. 8 aufzeigt, ist das Ruhephonemtakt-signal SPCL auf der Leitung 158 somit dem Phonemtakt-Aussignal PCO, das durch die Verzögerungsschaltung 162 verzögert und invertiert wird, äquivalent, und zwar ca. 10 μs .

Da der Flip-Flop 154 durch die ansteigende Flanke eines Signalspulses, den er an seinem Takteingang CL empfängt, verriegelt ist, wird ersichtlich, dass das Ruhephonemtakt-signal SPCL vor der Zeit t_1 zuerst den Flip-Flop 154 taktet, das Ruhephonemsignal SP aus dem Eingabebit 2⁸ bis dahin den D-Eingang des Flip-Flops 154 nicht erreicht hat. Deshalb hält das Schalttakt-signal auf der Leitung den $\bar{Q}$ -Ausgang des Flip-Flops trotzdem aufrecht und somit bleiben für einen Augenblick die Gatter 146 und 148 leitend. Wie das Diagramm aufzeigt, wird bis zur Zeit t_2 , wenn der Ausgang F der Logikschaltung 110

abfällt, die Anwesenheit eines Ruhephonems SP am D-Eingang des Flip-Flops 154 nicht festgestellt. Ca. 10 μs später, zur Zeit t_3 , fällt das Phonemtakt-signal aus dem Gatter 174 wieder ab, welches nach der weiteren Verzögerung von 10 μs , die durch die Schaltung 162 eingebracht werden, bewirkt, dass das Ruhephonemtakt-signal SPCL auf der Leitung 158 den Flip-Flop 154 wieder taktet. Somit wird zur Zeit t_4 etwa 300 μs nach dem ersten ansteigenden Impuls auf der Leitung 158, das Ruhephonemsignal SP aus dem Bit 2⁸ in den Flip-Flop 154 eingetragen. Dies lässt den $\bar{Q}$ -Ausgang des Flip-Flops 154 niedrig werden, welcher die Gatter 146 und 148 nichtleitend macht.

Diese durch die Schaltung 162 eingeführte weitere Verzögerung von 10 μs bildet eine Vorsichtsmassnahme, um sicherzustellen, dass das Ruhephonemsignal SP den D-Eingang des Flip-Flops 154 vor dem zweiten ansteigenden Impuls auf der Leitung 158 erreicht. Da das Ruhephonemsignal SP nicht vor der zweiten ansteigenden Flanke im Ruhephonemtakt-signal SPCL eingetragen wird, wird ein anderer ansteigender Impuls nicht auftreten, bis der nachfolgende Phonemtaktimpuls erzeugt wird, um das nächste Phonembefehlswort einzugeben. Das $\bar{Q}$ -Ausgangssignal aus dem Flip-Flop 154 bleibt somit für die Dauer des Phonemzeitraumes niedrig.

Um den Wechsel der logischen Zustände der Flip-Flops 120, 122, 124 und 126 zu verhindern, wenn ein Ruhephonem vorhanden ist, wird das Q-Ausgangssignal LSP aus dem Flip-Flop 154 an einen Eingang des NOR-Gatters 134 angelegt. Steigt das Q-Ausgangssignal LSP des Flip-Flops 154 zur Zeit t_4 an, so fällt der Ausgang des NOR-Gatters 134 ab, und zwar ohne Rücksicht des Zustandes des Signals FD' an seinen anderen Eingang. Dies hält daraufhin das Kennzeichentakt-signal FCL auf der Leitung 135 niedrig und verhindert die Verriegelung der Flip-Flops 120, 122, 124 und 126. Somit ist es wichtig, dass die durch die Schaltung 136 eingebrachte Verzögerung ausreichend lang ist, um sicherzustellen, dass das Ausgangssignal LSP des Flip-Flops 154 ansteigt (bei t_4) bevor das Signal FD' abfällt (bei t_5) deshalb liegt t_4 ca. 40 μs vor t_5 , deshalb kann die Gesamtsprechgeschwindigkeit und die Lautstärkenparameter des Audioausgangs während des Vorhandenseins eines Ruhephonems festeingestellt werden.



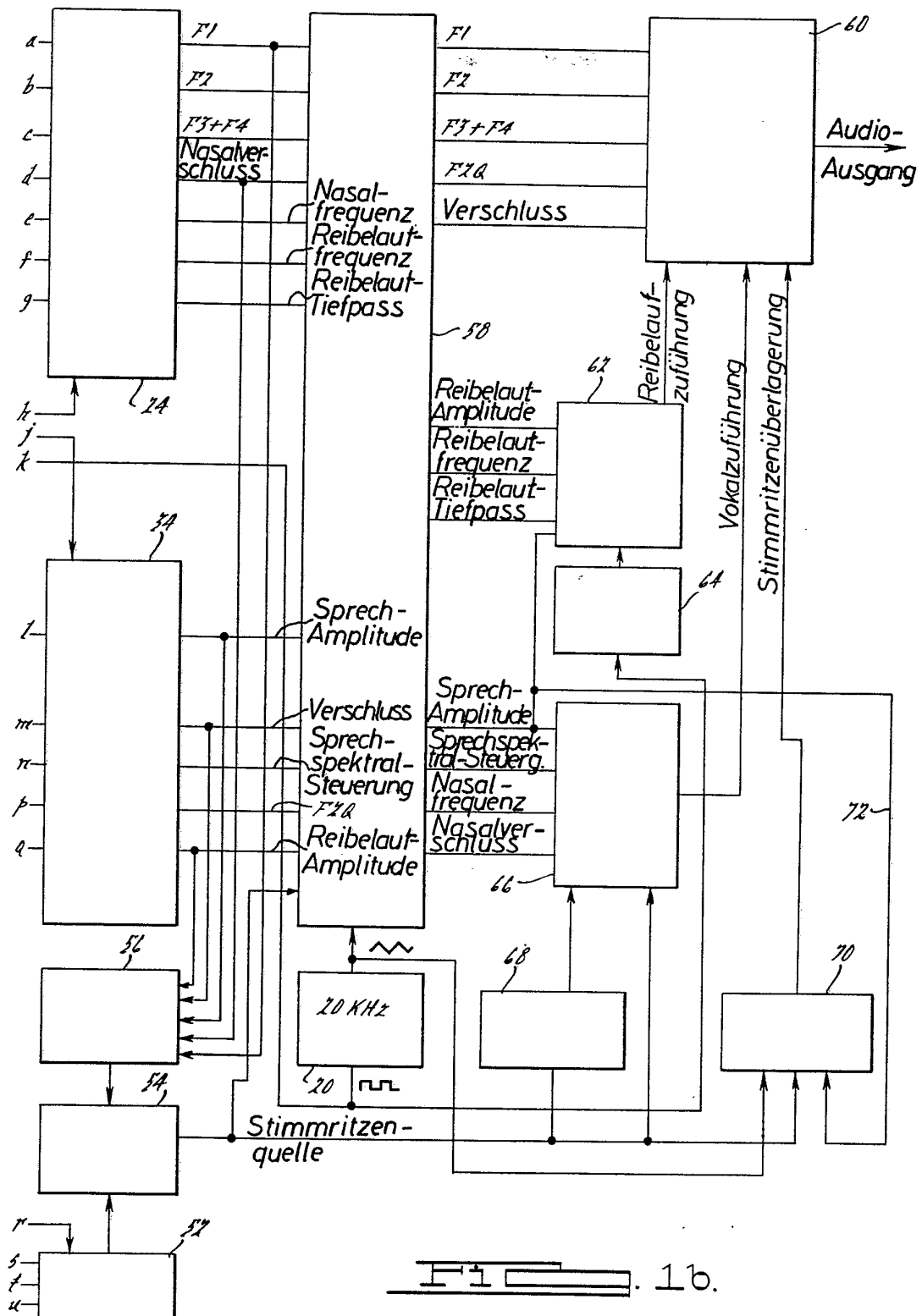
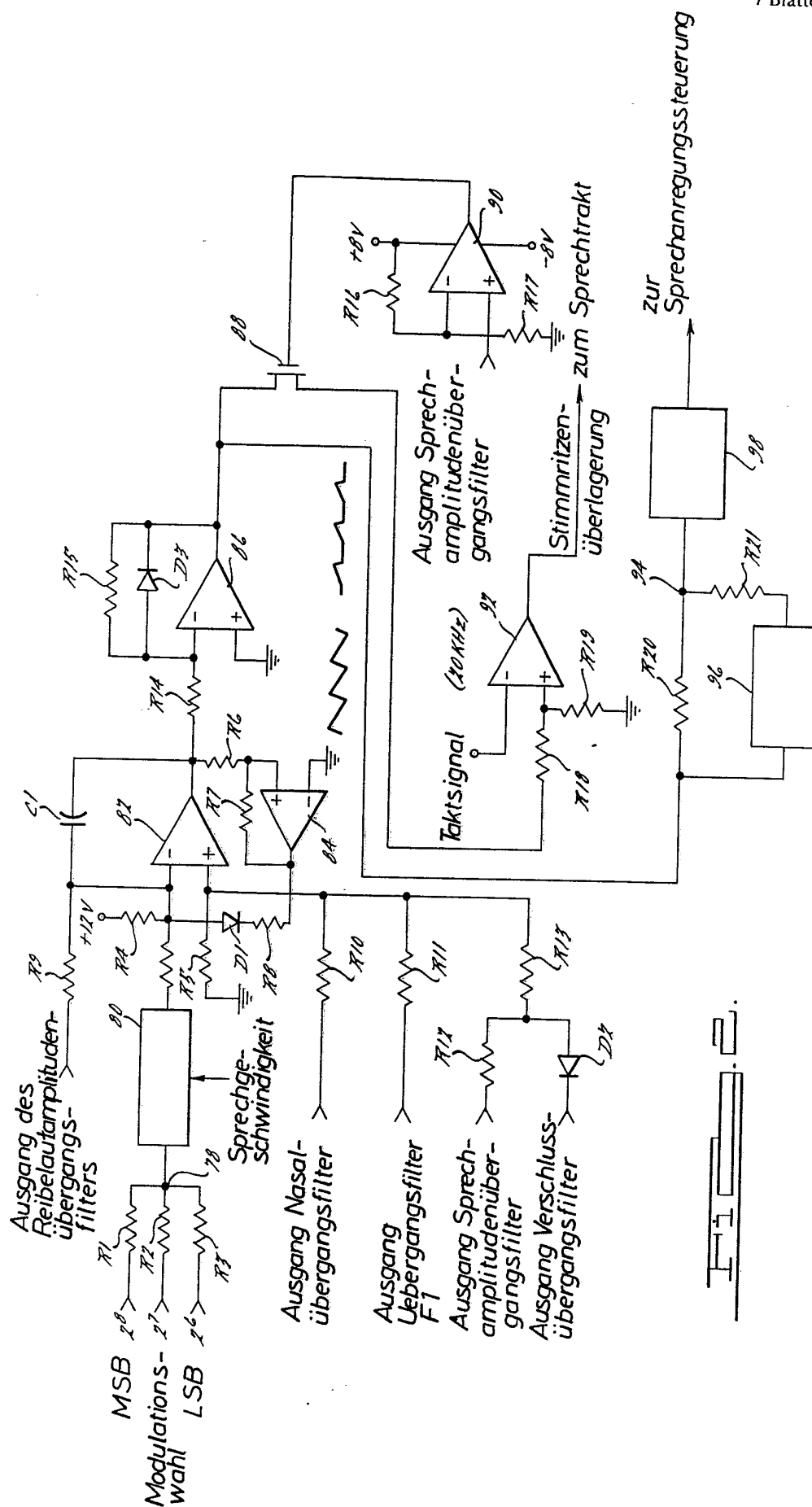
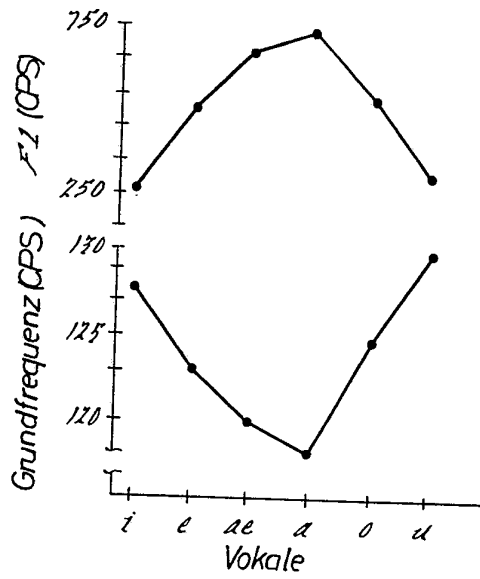


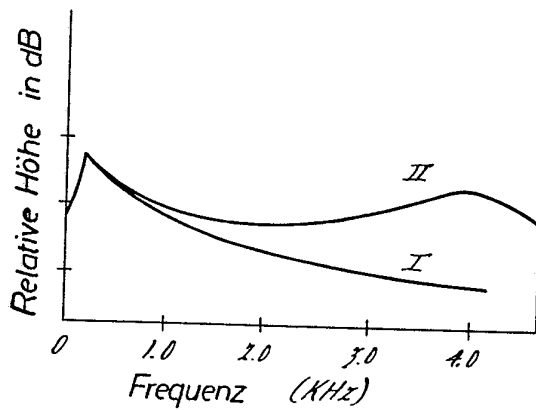
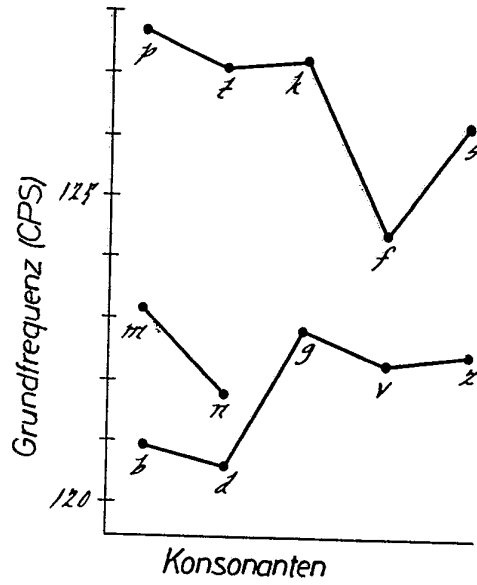
Fig. 1b.



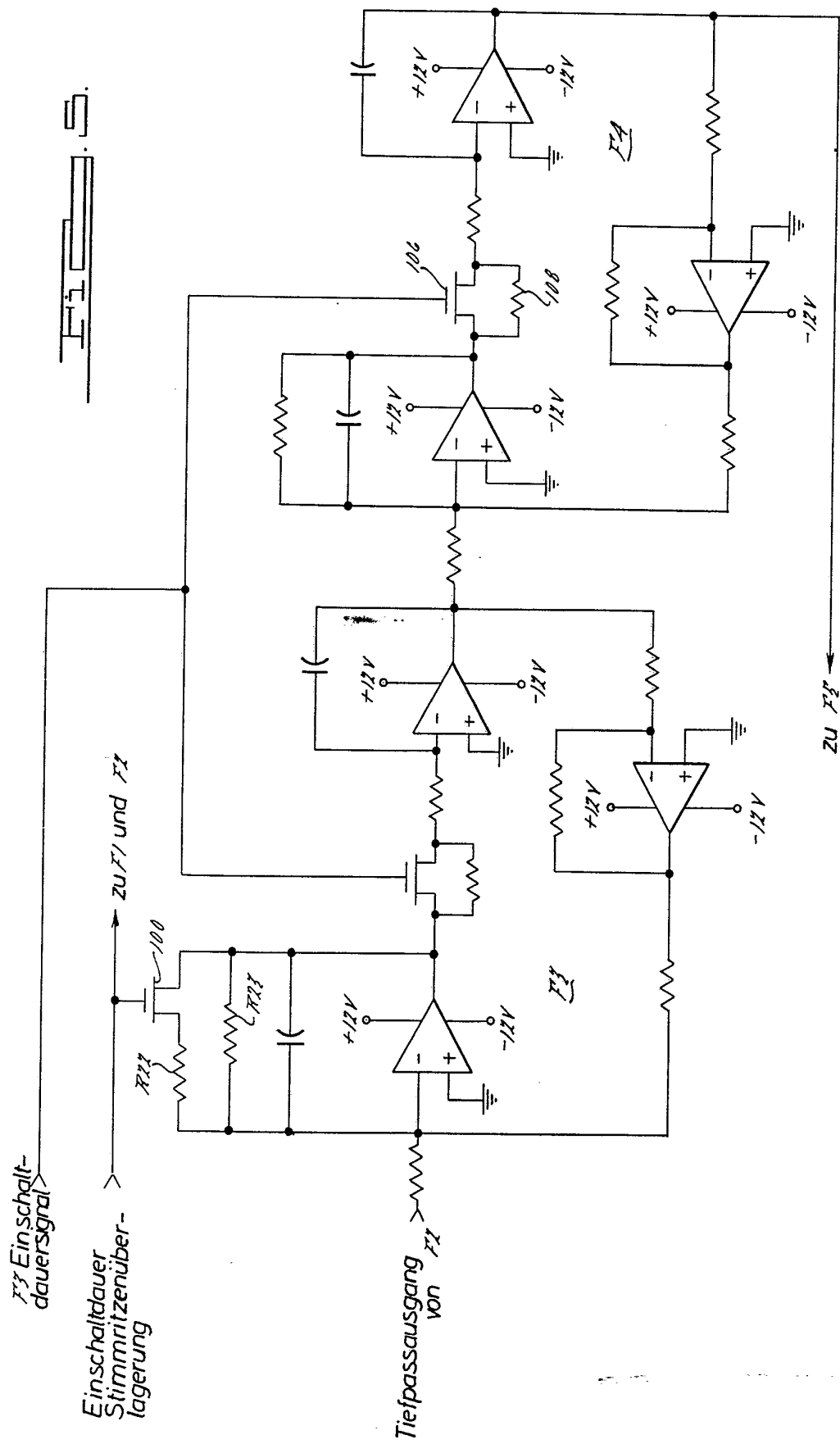


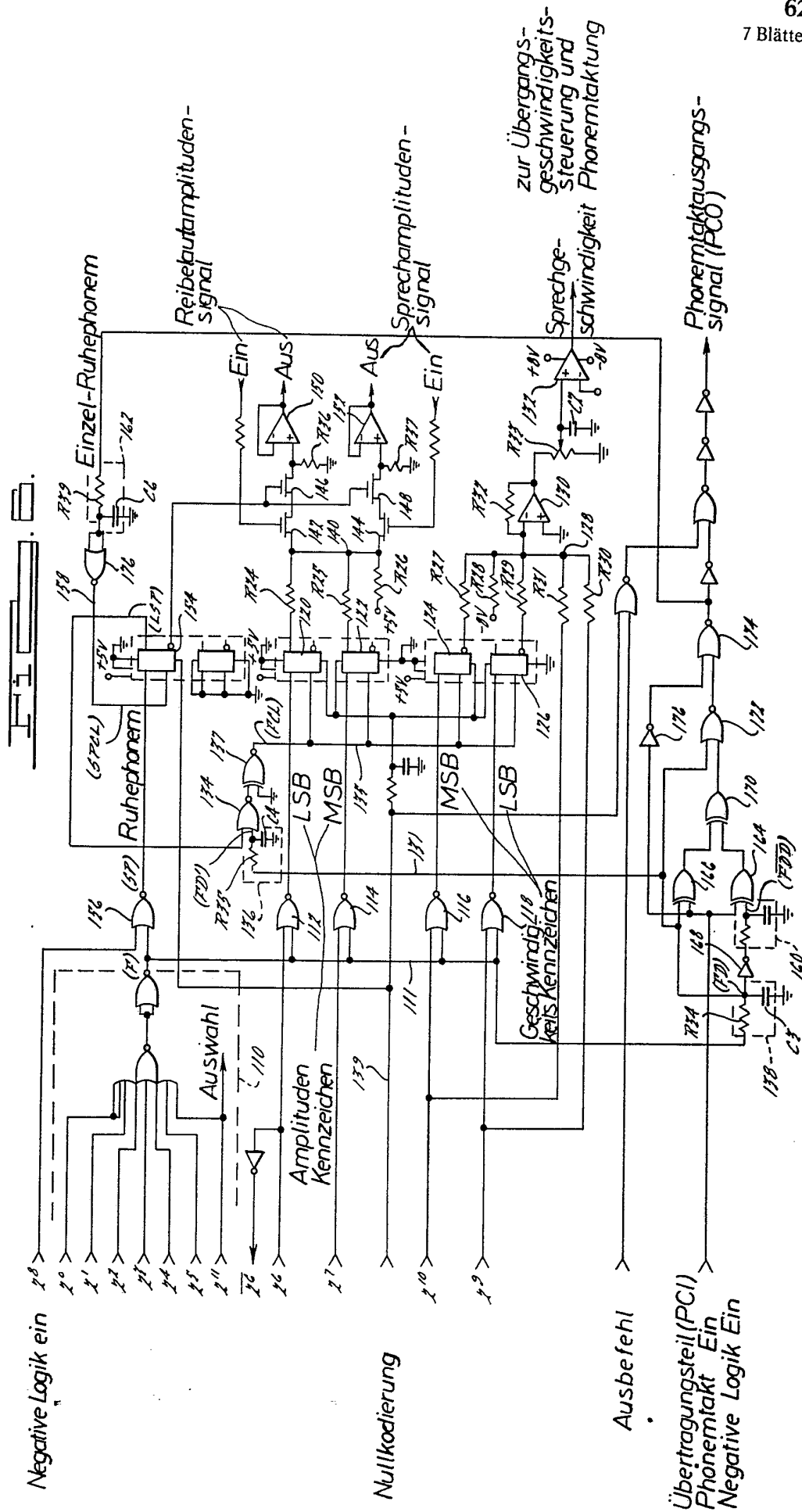
Fi. 3 a.

Fi. 3 b.



Fi. 4.





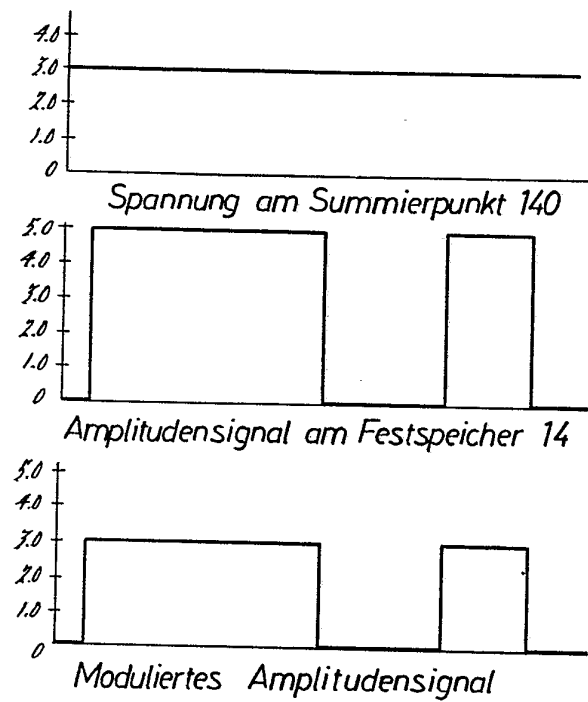


Fig. 7.

