



(11) **EP 4 047 595 A1**

(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:  
**24.08.2022 Patentblatt 2022/34**

(51) Internationale Patentklassifikation (IPC):  
**G10H 7/00<sup>(2006.01)</sup> G10H 7/04<sup>(2006.01)</sup>**

(21) Anmeldenummer: **22153060.3**

(52) Gemeinsame Patentklassifikation (CPC):  
**G10H 7/008; G10H 7/045**

(22) Anmeldetag: **24.01.2022**

(84) Benannte Vertragsstaaten:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**  
Benannte Erstreckungsstaaten:  
**BA ME**  
Benannte Validierungsstaaten:  
**KH MA MD TN**

(71) Anmelder: **e-instruments lab GmbH**  
**21079 Hamburg (DE)**

(72) Erfinder: **Paprotny, Alexander**  
**21614 Buxtehude (DE)**

(74) Vertreter: **Raffay & Fleck**  
**Patentanwälte**  
**Grosse Bleichen 8**  
**20354 Hamburg (DE)**

(30) Priorität: **23.02.2021 EP 21158720**

(54) **VERFAHREN ZUR ERSTELLUNG EINER GRUPPE VON PHASENKOHÄRENTEN SAMPLES FÜR EIN VIRTUELLES INSTRUMENT**

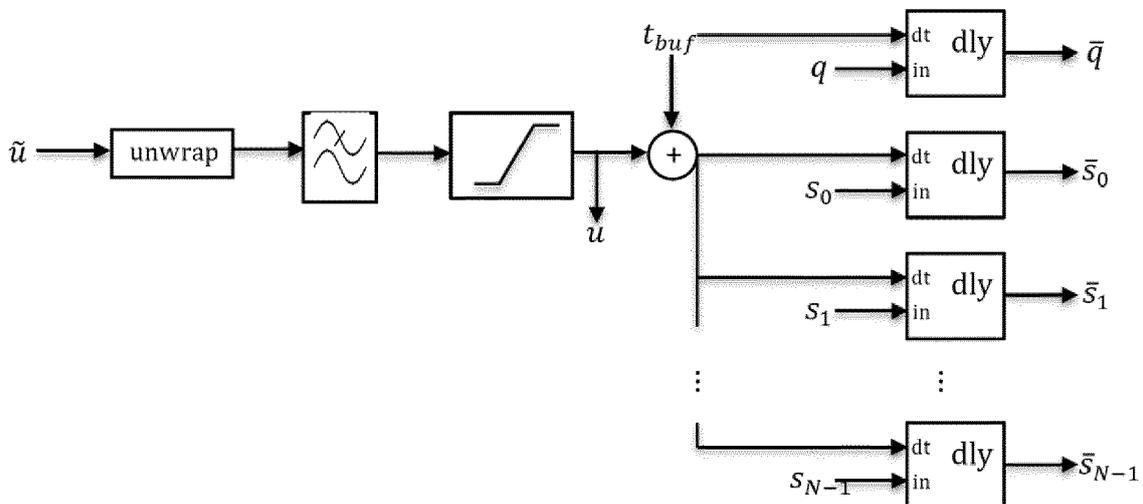
(57) Die vorliegende Erfindung betrifft phasenkohärente Samples, deren Nutzung in einem virtuellen Instrument, das virtuelle Instrument, und das Verfahren zur Erzeugung der Samples.

Bei virtuellen Musikinstrumenten auf Samplebasis besteht bei kontinuierlichen Übergängen zwischen Samples auf unterschiedlichen Dynamikstufen eines realen Instruments das Problem, dass es aufgrund von Phaseninkohärenz zwischen den Samples zu Kammfilterartefakten kommen kann.

Aufgabe ist es, diese Nachteile zu vermeiden und dennoch eine Lösung für kontinuierliche Übergänge zwischen Samples auf unterschiedlichen Dynamikstufen ei-

nes realen Instruments zu ermöglichen.

Gelöst wird die Aufgabe durch den Vergleich des Phasenverlaufs eines Referenzsamples mit dem Phasenverlauf mindestens eines weiteren der Samples über mindestens 500ms Sampledauer des Referenzsamples zur Erzeugung eines Verlaufes eines Phasenbeziehungswertes und Verändern der Geschwindigkeit und/oder Zeitbasis des mindestens einen weiteren auf Basis des Verlaufs des Phasenbeziehungswertes, wobei die Geschwindigkeit und/oder Zeitbasis über die Dauer von mindestens 500ms variierend abhängig vom Verlauf des Phasenbeziehungswertes verändert wird.



**Fig. 2**

**EP 4 047 595 A1**

## Beschreibung

**[0001]** Die vorliegende Erfindung betrifft phasenkohärente Samples, deren Nutzung in einem virtuellen Instrument, das virtuelle Instrument, und das Verfahren zur Erzeugung der Samples.

**[0002]** Das Abspielen mehrerer Samples in einem virtuellen Instrument ist beispielsweise aus der US 9, 805, 702 B1 bekannt.

**[0003]** Aus der US 2012/0243711 A1 ist ein Mixer mit mehreren Eingängen bekannt, auf die unterschiedliche aber statische Verzögerungen angewandt werden, um Laufzeitunterschiede zwischen Tonerzeuger zu den unterschiedlichen Mikrofonen zu kompensieren. Die Verzögerungen werden dabei automatisch bestimmt. Dazu wird ein einzelner Testton erzeugt, anhand dessen die Laufzeitunterschiede einmalig bestimmt und festgelegt werden.

**[0004]** Zudem ist es aus der US 2017/0372711 A1 ein (Re-)Syntheseverfahren, das also nicht mit Samples arbeitet. Dazu weist es einen Wellenformspeicher auf, in dem komprimiert Wellenformen gespeichert sind. Mehrere Tonerzeugungskanäle lesen jeweils eine Wellenform aus dem Speicher aus. Sofern ein von der gespeicherten Wellenform abweichender Pitch eingestellt ist, wird die Auslesegeschwindigkeit entsprechend angepasst bzw. werden durch lineare Interpolation entsprechende Zwischenwerte erzeugt. Dazu verwendet es für das Cross-fade zwischen mehreren Tonerzeugungskanäle einen Phasenzähler um die Dynamik des Cross-fades zu steuern.

**[0005]** Die JP 2001100756 A offenbart die einmalige Berechnung eines mittleren statischen Phasenversatzes über einen zeitlichen Teilabschnitt der Audiospuren auf Basis dessen ein einkanaliges Signal relativ zu einem einkanaligen Referenzsignal verschoben und dadurch das Referenzsignal zum einkanaligen Signal im Mittel phasenkohärenter ausgestaltet wird.

**[0006]** Aus der WO 2018/055892 A1 ist es zudem bekannt, ein Schlaginstrumentton zu synthetisieren. Dabei wird auf gespeicherte Wellenformen für unterschiedliche Schläge zurückgegriffen, die kombiniert werden. Dabei werden die Wellenformen jeweils in verschiedene Komponenten zerlegt gespeichert und zwar in einen Anteil, der Pitchhüllkurvendaten und Amplitudenhüllkurvendaten und Phasenstartwerte einer oder mehrerer Sinusanteile enthält und einen Restanteil. Die Daten der Sinusanteile der zu kombinierenden Töne werden bei einer Erstellung der für die Erzeugung zu verwendenden gewichtet gemischten Pitchhüllkurvendaten, Amplitudenhüllkurvendaten und Phasenstartwerte so verrechnet, dass eine negative Überlagerung ausgeschlossen ist. Sodann werden auf den gewichtet gemischten Pitchhüllkurvendaten, Amplitudenhüllkurvendaten und Phasenstartwerten basierend die Sinusanteile erzeugt und mit den gemischten Restanteilen vermischt. So kann eine auslöschende Überlagerung der Sinusanteile verhindert werden. Eine auslöschende Überlagerung nur der Restanteile wird vom Hörer weniger negativ empfunden.

**[0007]** Aus der DE 4008872 A1 ist ein Verfahren zur Speicherung und Synthese bekannt, bei dem Klänge aus einzelnen Klangmerkmalen zusammengesetzt werden, die in Form von Abtastwerten digital abgespeichert sind und die gesteuert ausgelesen werden, einer Abtastratenwandlung unterworfen werden, sodass alle Klangmerkmale mit einer einheitlichen Systemabtastrate zur Verfügung stehen, und dann zusammengesetzt werden. Dadurch lässt sich die Tonhöhe beim Abspielen variieren. Man nutzt dabei die Tatsache aus, dass bei den vielen Klangmustern, die in dem Speicher abgelegt sind, gewisse Gemeinsamkeiten vorherrschen, die beispielsweise das charakteristische Klangbild eines Instruments bestimmen, um Speicherplatz einzusparen. Das Mischen der Klangmerkmale erfolgt dann so, dass sie zum Beginn des Abspielens eine vordefinierte Phasenbeziehung aufweisen.

**[0008]** Bei virtuellen Musikinstrumenten auf Samplebasis besteht bei kontinuierlichen Übergängen zwischen Samples auf unterschiedlichen Dynamikstufen eines realen Instruments das Problem, dass es aufgrund von Phaseninkohärenz zwischen den Samples zu Kammfilterartefakten kommen oder psychoakustisch gar der Eindruck von zwei unisono spielenden Instrumenten entstehen kann. Derartige Pathologien wirken sich stets negativ auf den in der Regel angestrebten realistischen Klangeindruck aus. Zu Beseitigung der Artefakte ist eine dynamische Angleichung der Phasen der betroffenen Signale erforderlich, was technologisch jedoch eine erhebliche Herausforderung darstellt, wenn es sich um räumliche und/oder mehrkanalig aufgezeichnete Signale handelt.

**[0009]** Auf dem Markt für virtuelle Musikinstrumente sind zudem Lösungen zur Vermeidung oben geschilderter Probleme bekannt, die lediglich für anechoisch ("trocken") und einkanalig aufgezeichnete Signale geeignet sind. Auch ist es bekannt, für räumliche und mehrkanalige Signalen die Phasenkohärenz auf Kosten des originären Signalverlaufs herzustellen, indem die Signale modelliert und vollständig resynthetisiert werden. Dabei gehen natürliche Phasenschwankungen sowie zeitdynamische Veränderungen des Obertonspektrums verloren, was sich erheblich zu Lasten eines natürlichen, organischen und realistischen Klangeindrucks auswirkt.

**[0010]** Aufgabe ist es, diese Nachteile zu vermeiden und dennoch eine Lösung für kontinuierliche Übergänge zwischen Samples auf unterschiedlichen Dynamikstufen eines realen Instruments zu ermöglichen, ohne Kammfilterartefakten oder den psychoakustischen Eindruck von zwei unisono spielenden Instrumenten in Kauf zu nehmen und dabei natürliche Phasenschwankungen sowie zeitdynamische Veränderungen des Obertonspektrums zu erhalten.

**[0011]** Das hier vorgestellte Verfahren bietet eine Lösung für die genannten Probleme ohne Rückgriff auf Resynthesetechniken: Indem lediglich die

**[0012]** Abspielgeschwindigkeit der Originalsignale moduliert wird, bleiben natürliche Phasen- sowie Obertonspekt-

rumsschwankungen erhalten. Das Vorgehen ist demnach minimalinvasiv. Mehrkanalige und räumliche Signale können problemlos in eine phasenkohärente Form gebracht werden, wobei zusätzlich relative Phasenbeziehungen, etwa zwischen verschiedenen simultanen Mikrofonpositionen bei multimikrofonierten Aufnahmen, erhalten bleiben. Darüber hinaus kann das Verfahren auch auf einer gebrauchstüblichen sequenziellen Rechnerarchitektur als Echtzeitprozessor mit moderatem Rechenaufwand implementiert werden, was beispielsweise einen Einsatz als Plugin in einer digitalen Audioverarbeitungssoftware oder als Bordeffekt in einer Sampler-Engine ermöglicht.

**[0013]** Gelöst wird die Aufgabe insbesondere durch ein Verfahren zur Erstellung einer Gruppe phasenkohärenter Audioquellen, insbesondere Audiosignale oder Samples, umfassend:

- Bereitstellen mindestens jeweils einer Eingangsaudioquelle zweier Tonerereignisse, der gleichen Tonhöhe eines realen Musikinstruments,
- Auswählen eines Audiosignals, insbesondere eine der Eingangsaudioquellen, als Referenzsample,
- Vergleich des Phasenverlaufs des Referenzsamples mit dem Phasenverlauf mindestens einer (weiteren) der Eingangsaudioquellen über insbesondere mindestens 500ms Audioquellendauer des Referenzsamples zur Erzeugung eines Verlaufes eines über die Zeit variierenden Phasenbeziehungswertes,
- Verändern der Geschwindigkeit und/oder Zeitbasis der mindestens einen (weiteren) Eingangsaudioquelle über eine Dauer der weiteren Eingangsaudioquelle von insbesondere mindestens 500ms auf Basis des Verlaufs des Phasenbeziehungswertes, wobei die Geschwindigkeit und/oder Zeitbasis über die Dauer von insbesondere mindestens 500ms variierend abhängig vom Verlauf des Phasenbeziehungswertes verändert wird.

**[0014]** Dabei weist der Verlauf des Phasenbeziehungswertes insbesondere mindestens eine Änderung pro 100ms, insbesondere pro 10ms, auf und/oder wird die

**[0015]** Änderung der Geschwindigkeit und/oder Zeitbasis pro 100 ms, insbesondere pro 10ms, mindestens einmal geändert.

**[0016]** Unter Zeitbasis ist insbesondere die zeitliche Zuordnung der einzelnen Samplewerte des Samples und/oder Audioquelle zu verstehen.

**[0017]** Unter einer Eingangsaudioquelle wird insbesondere eine Audioquelle verstanden, die durch das Verfahren bearbeitet wird und/oder als Referenz für die Bearbeitung mindestens einer anderen Audioquellen durch das Verfahren verwendet wird.

**[0018]** Gelöst wird die Aufgabe auch durch ein Verfahren zum zeitgleichen Abspielen mindestens jeweils eines Samples zweier Tonerereignisse von mindestens 500ms Länge der gleichen Tonhöhe eines realen Instruments, wobei die Samples jeweils insbesondere über mindestens 500ms und zumindest teilweise zeitgleich phasenkohärent abgespielt werden.

**[0019]** Ebenfalls gelöst wird die Aufgabe durch ein virtuelles Instrument, insbesondere in VST-, AU- oder AAX-Plugin aufweisend mindestens eine Gruppe phasenkohärenter Samples und/oder eingerichtet zur Erzeugung solcher, wobei die Gruppe mindestens zwei Samples jeweils eines Tonerereignisse von insbesondere mindestens 500ms Länge der gleichen Tonhöhe eines realen Musikinstruments umfassend.

**[0020]** Dabei ist das virtuelle Instrument eingerichtet, mindestens zwei der phasenkohärenter Samples gleicher Tonhöhe zumindest teilweise gleichzeitig und phasenkohärent abzuspielen, wobei die Samples eine Länge insbesondere von mindestens 500ms aufweisen, in denen sie sich nicht wiederholen.

**[0021]** Dabei weist bevorzugt mindestens ein Sample der Gruppe eine über dessen Zeitverlauf variierende Geschwindigkeitsänderung auf und/oder ist mindestens ein Sample der Gruppe ein bearbeitetes Sampling, das ein RMS über seine zeitliche Länge, insbesondere ohne Ein- und Ausschwingzeiten, über die auf die mittlere Periode im Intervall normalisierten Abweichungen von der mittleren Periode des Intervalls und/oder von der Periode eines Referenzsamples aus der Gruppe von Samples von weniger als 75% des ebenso über den gleichen Zeitraum bestimmten RMS des unbearbeiteten Samples, aus dem das bearbeitete Sampling hervorgegangen ist und/oder des mittels eines statischen Zeitversatzes korrigierten Samples, aus dem das bearbeitete Sampling hervorgegangen ist, aufweist.

**[0022]** Dies trifft insbesondere für eine Mehrzahl der Samplings der Gruppe und/oder alle Samplings der Gruppe und/oder alle Samplings der Gruppe abgesehen vom Referenzsampling zu.

**[0023]** Die variierende Geschwindigkeitsänderung weist insbesondere mindestens eine Änderung pro 100ms, insbesondere pro 10 ms, auf.

**[0024]** Das virtuelle Instrument ist dabei insbesondere in Form einer Software und/oder Computerprogramm auf einem Datenträger gespeichert und/oder in Form eines Computers, der als virtuelles Instrument eingerichtet ist und/oder ein solches als Computerprogramm aufweist, ausgebildet.

**[0025]** Ebenfalls gelöst wird die Aufgabe durch ein System, insbesondere ein Hard- und Softwaresystem, aufweisend mindestens eine CPU und einen Speicher, eingerichtet zur Ausführung eines erfindungsgemäßen Verfahrens.

**[0026]** Gelöst wird die Aufgabe auch durch einen Signalprozessor mit mindestens zwei Eingängen für Audiosignale und mindestens zwei Ausgänge für Audiosignale und eingerichtet, die an den Audioeingängen anliegenden zu sampeln

und mit den erzeugten Samples das erfindungsgemäße Verfahren durchzuführen und die phasenkohärenten Samples an den Ausgängen auszugeben.

**[0027]** Gelöst wird die Aufgabe auch durch eine Gruppe von phasenkohärenter Samples unterschiedlicher Dynamiken einer Tonhöhe eines realen Musikinstruments, wobei die Samples eine Länge von insbesondere mindestens 500ms aufweisen, in denen sie phasenkohärent sind und sich nicht wiederholen.

**[0028]** Dabei weist bevorzugt mindestens ein Sample der Gruppe eine über dessen Zeitverlauf variierende Geschwindigkeitsänderung auf und/oder ist mindestens ein Sample der Gruppe ein bearbeitetes Sampling, das ein über seine zeitliche Länge, insbesondere ohne Ein- und Ausschwingzeiten, bestimmtes RMS über die auf die mittlere Periode im Intervall normalisierten Abweichungen von der mittleren Periode des Intervalls und/oder von der Periode eines Referenzsamples aus der Gruppe von Samples von weniger als 75% des ebenso über den gleichen Zeitraum bestimmten RMS des unbearbeiteten Samples, aus dem das bearbeitete Sampling hervorgegangen ist und/oder des mittels eines statischen Zeitversatzes korrigierten Samples, aus dem das bearbeitete Sampling hervorgegangen ist, aufweist.

**[0029]** Dies trifft insbesondere für eine Mehrzahl der Samplings der Gruppe und/oder all Samplings der Gruppe und/oder alle Samplings der Gruppe abgesehen vom Referenzsampling zu.

**[0030]** Die variierende Geschwindigkeitsänderung weist insbesondere mindestens eine Änderung pro 100ms, insbesondere pro 10ms, auf.

**[0031]** Gelöst wird die Aufgabe auch durch eine Gruppe von Samples umfassend Samples unterschiedlicher Dynamiken einer Tonhöhe eines realen Musikinstruments, wobei die Samples jeweils eine Länge von mindestens 500ms aufweisen, und wobei die Gruppe Informationen zum Phasenverlauf mindestens eines, insbesondere aller Samples, über die Länge von mindestens 500ms der Samples aufweist, insbesondere als Spur, insbesondere Midi-Spur, insbesondere jeweils einer zusätzlichen Spur die Information zum Phasenverlauf beinhaltend, insbesondere einer je Sample. Eine solche Gruppe kann verwendet werden, um auf einfache Weise ein phasenkohärentes Abspielen zu erreichen, indem beim Abspiel mindestens zweier der Samples, zumindest teilweise zeitgleich, die Unterschiede des Phasenverlaufs der Samples während der zeitgleichen Wiedergabe, durch Modulation der Abspielgeschwindigkeit(en) der/des Samples verringert werden. Dabei kann der Verlauf des Phasenunterschiedes auf einfache Weise aus der Information zum Phasenverlauf gewonnen werden. Dies bietet unter anderem den Vorteil, dass der zeitliche Versatz relativ frei wählbar ist und/oder Glättungs- und/oder Vorausschauparameter durch den Benutzer einstellbar ausgestaltet sein können.

**[0032]** Auch die Verwendung einer Gruppe phasenkohärenter Audioquellen, insbesondere Audiosignalen oder Samples, mindestens zweier unterschiedlicher Dynamiken einer Tonhöhe eines realen Musikinstruments löst die Aufgabe, wobei die Audioquellen eine Länge von mindestens 500ms aufweisen, in denen sie phasenkohärent sind und sich nicht wiederholen, zur phasenkohärenten und zumindest teilweise gleichzeitigen Wiedergabe der Audioquellen einer Tonhöhe mit mindestens zwei unterschiedlichen Dynamiken. Auch die Verwendung einer Gruppe von Audioquellen, insbesondere Audiosignalen oder Samples, mindestens zweier unterschiedlicher Dynamiken einer Tonhöhe eines realen Musikinstruments löst die Aufgabe, wobei die Audioquellen eine Länge von mindestens 500ms aufweisen, in denen sie Informationen zum Phasenverlauf aufweisen, zur phasenkohärenten und zumindest teilweise gleichzeitigen Wiedergabe der Audioquellen einer Tonhöhe mit mindestens zwei unterschiedlichen Dynamiken. Dies erfolgt insbesondere durch Ermitteln eines Verlaufes eines Phasenbeziehungswerts, insbesondere Ermitteln eines Verlaufs der Phasenabweichung, über die Dauer der Audioquelle für jedes der abhängigen Audioquellen und Verwenden des jeweils ermittelten Verlaufs zur Anpassung der Geschwindigkeit und/oder Zeitbasis der jeweiligen abhängigen Audioquelle. Der Phasenbeziehungswert verändert sich insbesondere über die Dauer des Samples.

**[0033]** Das besondere des erfindungsgemäßen Verfahrens liegt insbesondere darin, eine dynamische Anpassung zu realisieren, durch die auch bei sich im Zeitverlauf ändernden Phasenbeziehungen zwischen den Eingangssignalen eine dauernde Phasenkohärenz erreichbar ist und bei der es zudem möglich ist zahlreiche Aufnahmen des gleichen Tonerignisses phasenkohärent zu gestalten.

**[0034]** Der Vergleich des Phasenverlaufs kann allgemein beispielsweise mittels eines Kreuzkorrelators erfolgen. Bevorzugt wird es aber, Flanken der Audiosignale zu bestimmen und basierend auf der zeitlichen Lage der Flanken in den Samples einen Vergleich des Phasenverlaufs durchzuführen. Dazu können beispielsweise die Durchgänge des Signals, insbesondere der Grundfrequenz, durch einen vorgegebenen Pegel, insbesondere die Nulldurchgänge bestimmt werden und die Lage der Nulldurchgänge als Information über die zeitliche Lage der Flanken verwendet werden. Insbesondere erfolgt die Anpassung so, dass der Versatz der Flanken, insbesondere Nulldurchgänge, reduziert wird.

**[0035]** Aus der Lage der Flanken zweier Eingangsaudioquelle, insbesondere des Referenzsamples und einer abhängigen/weiteren Audioquelle wird insbesondere der Verlauf eines Phasenvergleichswerts ermittelt, der insbesondere die Abweichung der Lage der Flanken angibt.

**[0036]** Der Verlauf eines Phasenvergleichswerts wird insbesondere geglättet. Dies erfolgt insbesondere durch einen Tiefpassfilter, dessen Eckfrequenz insbesondere so eingestellt wird, dass bei Anwendung zur Modulation der Samples und zeitgleichen Abspielen der Samples keine Artefakte, insbesondere keine Glitches, entstehen. Um die durch den Tiefpassfilter entstehende Zeitverzögerung zumindest teilweise auszugleichen, kann bei der Bestimmung des Verlaufs

des Phasenvergleichswertes eine Vorausschau (look ahead) verwendet werden, der insbesondere nach Einstellung des Tiefpassfilters eingestellt wird und/oder so eingestellt wird, insbesondere so lange erhöht wird, dass/bis keine Kammfilterartefakte auftreten. Die Vorausschau liegt bevorzugt im Bereich von 1 bis 50 ms, insbesondere über 5 ms und/oder unter 20 ms. Dadurch entsteht insbesondere der Verlauf eines Phasenbeziehungswerts, auch Steuerfunktion, der bevorzugt zur Modulation der Abspielgeschwindigkeit und/oder zum Verändern der Geschwindigkeit und/oder Zeitbasis verwendet wird. Somit wird es bevorzugt, zur Modulation der Abspielgeschwindigkeit und/oder zum Verändern der Geschwindigkeit und/oder Zeitbasis einen Mittelwert zu verwenden, der den Phasenverlauf des Referenzsamples und des abhängigen Samples, das verändert wird, vor, an und nach dem Zeitpunkt, an dem jeweils verändert wird, berücksichtigt.

**[0037]** Der Glättungsgrad wird insbesondere mittels eines Reglers über die Eckfrequenz eines auf das Steuerungssignal angewandten Tiefpassfilters eingestellt. Eine hohe Eckfrequenz hat einen niedrigen Glättungsgrad zu Folge und vice versa. Allgemein gilt hier, dass mit einem zunehmenden Glättungsgrad messfehler- oder signalqualitätsbedingte Artefakte ("glitches") vermieden werden, jedoch auch durch die entsprechend erhöhte Reaktionszeit der Wirkungsgrad des Verfahrens gemindert wird und demzufolge phasenverschiebungsbedingte Kammfiltereffekte hörbar werden können. Durch auf geeignete Einstellung des Vorausschau-Parameters ("look ahead") kann auch bei hohen Glättungsgraden eine wirksame Reaktion des Schaltkreises auf sprunghafte Änderungen des Phasendifferenzverlaufs erzielt werden. Allerdings sollte hierbei beachtet werden, dass bei hohen Vorausschauzeiten und geringer Glättung der Schaltkreis verfrüht reagiert und somit unter Umständen selbst Kammfiltereffekte erzeugen kann. Der Parameter wird insbesondere auf einen reellen Wert zwischen null und eins eingestellt, wobei bei letzterer Einstellung die gesamte Pufferzeit (üblicher Weise bei ca. 60ms) zur Vorausschau genutzt wird (d.h., dass Steuerungssignal wird um die Pufferzeit verfrüht angewandt bzw. die zu steuernden Signale um die Pufferzeit verzögert), bei ersterem (Wert von 0) hingegen überhaupt keine Vorausschau stattfindet. Die beiden Parameter werden bevorzugt durch den Benutzer signalspezifisch nach Gehör so eingestellt, dass sich in Hinblick auf das gewünschte Ergebnis ein geeigneter Kompromiss zwischen geringer Reaktionszeit und Artefaktfreiheit ergibt. Mit einer solchen Vorgehensweise ist ein Toningenieur gemeinhin vertraut, da eine Vielzahl branchenüblicher Effektgeräte auf eine ähnliche Weise bedient wird. Etwa werden bei einem Dynamikkompressor Reaktionszeit ("attack" und "release") sowie Vorausschauzeit nach Gehör auf eine solche Weise eingestellt, dass sich ein zufriedenstellender Kompromiss zwischen Dynamikreduktion und Artefaktfreiheit ergibt.

**[0038]** Bevorzugt berücksichtigt der Wert des Verlaufs des Phasenbeziehungswertes an einem Zeitpunkt des Referenzsamples zumindest teilweise den Phasenverlauf des Referenzsamples und/oder den Phasenverlauf der einer weiteren der Audioquellen, vor, an und/oder nach dem Zeitpunkt des Referenzsamples. Dabei wird die Zeit im Referenzsample und in der weiteren Audioquelle insbesondere identisch, insbesondere jeweils von deren Beginn und/oder vorbestimmten Zeiten gezählt. Eine solche Berücksichtigung erfolgt insbesondere durch eine Vorausschau und/oder einen Tiefpass und/oder eine Mittelwertbildung. Insbesondere erfolgt die Berücksichtigung über den gesamten Verlauf, mit Ausnahme des Beginns und Endes, insbesondere über jeweils 1 bis 100 ms zu Beginn und zum Ende der weiteren Audioquelle und/oder des Referenzsamples.

**[0039]** Eine Audioquelle kann dabei mehrere Spuren enthalten, beispielsweise einen oder mehrere Stereokanäle. Jede Audioquelle ist insbesondere die Aufnahme eines realen Tonereignisses, beispielsweise mittels eines Mikrofons oder einer Mehrzahl von Mikrofonen, insbesondere an einer gemeinsamen Position. Statische Phasenverschiebungen auf Grund unterschiedlicher Laufzeiten zu unterschiedliche Mikrofonposition, insbesondere unterschiedlicher Audioquellen, eines gemeinsamen Tonereignisses stellen keinen Nachteil dar. Sie werden durch das Verfahren insbesondere erhalten und nicht beseitigt.

**[0040]** Eine Audioquelle ist dabei insbesondere die Aufnahme eines Tonereignisses, insbesondere an einer Position, erzeugt durch das Anschlagen oder Anstreichen einer Saite oder das Spielen eines Tons auf einem Blasinstrument.

**[0041]** Die so veränderten abhängigen Audioquellen bilden alleine und/oder zusammen mit dem Referenzsample eine Gruppe kohärenter Audioquellen. Als Referenzsample kann eine der Audioquellen genutzt oder ein Audiosignal, insbesondere mit konstanter Frequenz und insbesondere ohne Phasensprung, insbesondere ein Sinussignal, genutzt werden. Letzteres ermöglicht das zeitversetzte phasenkohärente Abspielen ohne eine an den Zeitversatz angepasste Phasenkorrektur vorzunehmen, beseitigt aber die eigentlich gewünschten zeitlichen Schwankungen der der Frequenz der Samples oder der phasenkohärent wiedergegebenen Audioquellen.

**[0042]** Zur Erzeugung kohärenter Audioquellen wird insbesondere eine Gruppe oder Mehrzahl von Audioquellen vorgehalten, die nach dem erfindungsgemäßen Verfahren in kohärente umgewandelt werden können. Diese liegen insbesondere auf einem Datenträger gespeichert vor und/oder sind insbesondere Teil des Systems und/oder virtuellen Instrument.

**[0043]** Vorteilhafterweise wird ein erster Zeitabschnitt, insbesondere die Dauer des Einschwingvorgangs, mit einer Länge insbesondere im Bereich von 0 bis 500 ms, insbesondere im Bereich von 100 bis 400 ms, der abhängigen Audioquellen in der Geschwindigkeit und/oder Zeitbasis nicht verändert und ein zweiter Zeitabschnitt, insbesondere mit einer Länge von mehr als 500 ms und/oder unmittelbar folgend auf den ersten Zeitabschnitt, der oder jeder der abhängigen Audioquellen in der Geschwindigkeit und/oder Zeitbasis verändert.

**[0044]** Der erste und/oder zweite Zeitabschnitt ist dabei insbesondere eine Aufnahme eines realen Tonereignisses, die nicht wiederholt wird/ist und/oder sich nicht wiederholt.

**[0045]** Die Gruppe von kohärenten Audioquellen ist insbesondere so ausgebildet, dass sie bei zeitgleichem Start kohärent sind, insbesondere zumindest über eine Dauer von mindestens 500ms, in der insbesondere keine Wiederholung vorliegt und/oder zumindest nach einem/dem ersten Zeitabschnitt.

**[0046]** Die Gruppe von kohärenten Audioquellen ist insbesondere so ausgebildet, dass sie eine Information aufweist oder von einer solchen begleitet ist, die angibt, in welcher zeitlichen Abfolge die teilweise zeitgleiche Wiedergabe zu starten ist, damit die Wiedergabe im Bereich der zeitgleichen Wiedergabe kohärent ist.

**[0047]** Anstelle und/oder zusätzlich zum Vorhalten einer Gruppe kohärenter Audioquellen kann die Gruppe kohärenter Audioquellen auch abhängig von Einstellungen und/oder dem zeitlichen Versatz der zumindest teilweise zeitgleichen Wiedergabe abhängig von diesen/m bei Bedarf, insbesondere vor (insbesondere unmittelbar vor und/oder maximal eine Stunde) und/oder während der Wiedergabe, erzeugt werden. Dies ermöglicht mehr Freiheiten in Bezug auf den zeitlichen Versatz der Wiedergabe bzw. deren Start.

**[0048]** Das Verfahren zum Abspielen, die Verwendung, das System oder das virtuelle Instrument sind insbesondere so ausgebildet, dass sie Lautstärke der zumindest teilweise zeitgleich wiedergegebenen Samples relativ zueinander veränderbar ist, insbesondere über die Dauer der Wiedergabe variierend. Dazu sind insbesondere entsprechende Einstellmöglichkeiten vorgesehen.

**[0049]** Der zeitliche Phasendifferenzverlauf zum entsprechenden Mikrofonsignal einer anzulegenden Dynamikstufe wird bevorzugt durch einen Phasenkorrelator (nicht notwendigerweise ein Kreuzkorrelator) ermittelt.

**[0050]** Nachdem ein Verlauf eines Phasenvergleichswerts erzeugt wurde, wird dieser bevorzugt durch geeignete Entklappung, Clipper, Filter und/oder Dynamikprozessoren verarbeitet und/oder geglättet, beispielsweise mittels Tiefpasses mit Eck- oder Grenzfrequenz im Bereich von 50 bis 200 Hz, insbesondere von bis 60 bis 150Hz.

**[0051]** Ein Entklappungsverfahren erkennt die Sprungstellen und rekonstruiert auf dieser Grundlage den stetigen Phasendifferenzverlauf. Dies kann beispielsweise erfolgen, wie in "One-Dimensional Phase Unwrapping Problem" Dr. Munther Gdeisat and Dr. Francis Lilley z.B. unter [https://www.ljmu.ac.uk/~media/files/ljmu/about-us/faculties-and-schools/fet/geri/onedimensionalphaseunwrapping\\_finalpdf.pdf](https://www.ljmu.ac.uk/~media/files/ljmu/about-us/faculties-and-schools/fet/geri/onedimensionalphaseunwrapping_finalpdf.pdf) beschrieben. Es kann aber alternativ oder zusätzlich die Abfolge der erkannten Nulldurchgänge verwendet werden, da in bei fehlerfreier Bestimmung der Flanken ein Umklappen der Phasendifferenzfunktion unmittelbar aus dem Vorliegen von zwei aufeinander folgenden Impulsen aus einem der Signale ohne einen zwischenzeitlichen Impuls des anderen Signals ersichtlich wird. Um nun wiederum durch Flankenmessfehler bedingte Artefakte und Sprünge zu vermeiden, kann ein bekanntes Entklappungsverfahren mit dieser Methode kombiniert werden. Der Entklappungsvorgang wird dadurch präziser und weniger anfällig für Störungen, was wiederum in vielen Fällen eine weniger invasive Glättung und somit eine etwas kürzere Reaktionszeit ermöglichen kann.

**[0052]** Ein Clipper kann dazu dienen, Signalspitzen jenseits eines vorgegebenen Frequenz-Intervalls abzuschneiden. Die untere Intervallgrenze des Clippers ist bevorzugt identisch mit dem Inversen der Pufferzeit, mit der das Abspielen der Samples Grundverzögert wird, die obere bevorzugt fest eingestellt, beispielsweise auf einen Wert im Bereich von  $10^{(-7)}$  bis  $10^{(-9)}$  Hz eingestellt.

**[0053]** Dadurch wird bevorzugt der Verlauf eines Phasenbeziehungswerts, auch Steuerfunktion, erzeugt.

**[0054]** Der Verlauf eines Phasenbeziehungswerts wird insbesondere als Zeitparameter in Verzögerungsglieder eingespeist, an deren Eingangsbusse jeweils ein abhängiges Samples angelegt wird. An den Ausgängen der Verzögerungsglieder können nun die zum Quellsignal phasenkohärenten, modulierten Audioquellen abgegriffen werden.

**[0055]** Bei mehrkanaligen und/oder mehrspurigen Samples kann der Verlauf eines Phasenbeziehungswerts beispielsweise nur für einen Kanal/eine Spur, für mehrere Kanäle/Spuren getrennt oder mittels einer mehrere Kanäle/Spuren berücksichtigender Metrik ermittelt werden. Die Metrik kann beispielsweise ein Mittelwert sein.

**[0056]** Das Verändern der Zeitbasis, kann beispielsweise für jeden Kanal/jede Spur eines abhängigen Samples getrennt und/oder unterschiedlich und/oder für alle Kanäle/alle Spuren eines abhängigen Samples gemeinsam und/oder gleich erfolgen.

**[0057]** Mit Vorteil weisen alle Audioquellen und/oder alle abhängigen Audioquellen die gleiche Spur- und/oder Kanalanzahl auf. Das Verfahren lässt sich aber auch anwenden, wenn die Spur- und/oder Kanalanzahl unterschiedlich ist.

**[0058]** Mit Vorteil weisen alle Audioquellen und/oder alle abhängigen Audioquellen die gleiche Samplingrate auf. Bei den Audioquellen handelt es sich insbesondere um analog oder digitale Aufnahmen analoger akustischer Ereignisse.

**[0059]** Mit Vorteil weisen alle Audioquellen und/oder alle abhängige Audioquellen die gleiche Dauer auf. Insbesondere weist das Referenzsample mindestens die Dauer der zweitlängsten, insbesondere der längsten, abhängigen Audioquellen auf.

**[0060]** Bei stark räumlichen Signalen kann es dazu kommen, dass beispielsweise aufgrund von diffus-schallbedingten Verzerrungen die Phasenbeziehungen zwischen den Signalen nicht korrekt durch den Korrelator erkannt werden. So kann es bevorzugt sein, dass Referenzsamples durch ein näher als alle/die Mikrofone der abhängigen Audioquellen, insbesondere sehr nahe an der Schallquelle, insbesondere maximal 1 Meter von der Schallquelle entfernt, positioniertes Mikrophon aufzunehmen. Dieses Referenzsamples ist dann bevorzugt nicht Teil der Gruppe kohärenter Audioquellen.

**[0061]** Bevorzugt sind die Audioquellen Aufnahmen eines zu dem/den verwendeten Mikrofon(en) ortsfesten Tonerignis. Das bedeutet, das bevorzugt Musikinstrument und Mikrofone zur Erstellung der Audioquellen ortsfest sind.

**[0062]** Vorteilhaft ist es zum einen, ein echtzeitmäßig reagierendes Phasenbeziehungsbestimmungsmodul zur Ermittlung des Verlaufs der Phasenabweichung zu verwenden, damit die für monophone Musikinstrumentensamples charakteristischen kurzzeitig auftretenden Phasenschwankungen und -abweichungen rasch ausgeglichen werden können.

**[0063]** Die Bestimmungsdauer für das zeitliche Intervall zwischen zwei Flanken ist bevorzugt im Wesentlichen durch dieses gegeben ist, da bei Ankunft einer Flanke ja sozusagen auf die korrespondierende Flanke des zweiten Signals gewartet werden muss. Die Gesamtreaktionszeit ist also in sehr guter Näherung gegeben durch die absolute zeitliche Phasendifferenz der gegebenen Signale plus einen Wert, der auf der Glättung, insbesondere deren Eck- oder Grenzfrequenz basierend, der bevorzugt im Bereich von 1 bis 20 ms liegt. Somit liegt die Reaktionszeit bevorzugt unter 100ms, insbesondere unter 50ms.

**[0064]** Zum anderen ist zur Vermeidung von messungs- oder signalbedingten Artefakten (etwa bewirkt durch ausfallende Fundamentaltöne (häufig in den tiefen Lagen tieftöniger (Tenor- und Bass-)Instrumente, wie etwa einem Cello oder Kontrabass) eine geeignete Glättung und/oder Siebung des Verlaufs der Phasenabweichung vor Nutzung zur Veränderung der abhängigen Samples vorteilhaft, beispielsweise mittels Tiefpass mit Eck- oder Grenzfrequenz im Bereich von 50 bis 200 HZ, insbesondere von bis 60 bis 150Hz.

**[0065]** Alternativ oder zusätzlich kann zur Vermeidung von messungs- oder signalbedingten Artefakten, insbesondere bei (kurzzeitig) ausfallendem/n Fundamentaltöne(n), vorübergehend oder dauerhaft anstelle des Grundtons die erste Harmonische des Grundtons als Referenz verwendet wird. Nun kann dies jedoch aufgrund von Phasenauslöschungen, die wiederum durch dem Umstand bedingt sind, dass die erste Harmonische prinzipiell keinerlei Informationen bezüglich der Richtung der korrespondierenden Flanken in der Fundamentalen enthält, zu neuerlichen Artefakten, insbesondere nunmehr künstlich hervorgerufenen Fundamentaltönen bei synchronem Abspielen der resultierenden phasenangeglichenen Samples kommen. Dies kann jedoch wiederum auf einfache Weise durch eine statische Polaritätsumkehr des prozessierten Referenzsignals behoben werden. Bei der Polaritätsumkehr (oftmals sachlich nicht gänzlich korrekt als "Phasendrehung" bezeichnet) handelt es sich ebenfalls um einen standardmäßigen Vorgang, welcher in marktüblichen Mischpulten und digitalen Audioverarbeitungsumgebungen implementiert ist. Alternativ kann auch das prozessierte Referenzsignal um eine halbe Periode verschoben werden.

**[0066]** Die Ermittlung des Verlaufs eines Phasenbeziehungswerts kann auf verschiedene Arten, beispielsweise analog, digital, in Echtzeit oder zeitlich nachgelagert, erfolgen. Bevorzugt wird zur Ermittlung des Verlaufs eines Phasenbeziehungswerts zunächst mit einem steilflankigen Bandpassfilter der Grundton des Referenzsamples und/oder der Audioquellen isoliert.

**[0067]** Um den Phasenversatz zwischen Referenzsample und abhängigen Audioquellen zu ermitteln, werden bevorzugt die sinusförmigen Grundtonverläufe durch einen Komparator prozessiert, woraus eine Rechtecksignal resultiert, welche sodann differenziert werden kann. Hierbei entsteht jeweils eine Folge von Diracimpulsen, welche sich gerade an den Nulldurchgängen der Samples befinden und je nach Richtung des Nulldurchgangs positiv bzw. negativ signiert sind. Es können aber auch sowohl positiv, als auch negativ signierte Flanken zur Ermittlung des Pha-Phasenversatz berücksichtigt werden. Dies führt zu einer Verdoppelung der Auflösung der Phasendifferenzermittlung und kann somit zu kürzeren Reaktionszeiten, beispielsweise innerhalb einer halben Periode des Signals, führen.

**[0068]** Bevorzugt wird der zeitliche Abstand zwischen Nulldurchgängen des Referenzsamples und je eine abhängige Audioquelle bestimmt und/oder verwendet. Die daraus bevorzugt erzeugte Treppenfunktion zeigt nach Entklappung ("unwrapping") den Phasendifferenzverlauf zwischen den beiden Samples an.

**[0069]** Zwecks Kompensation von Messfehlern und Signalaussetzern, sowie zur Vermeidung von Artefakten, die durch zu große Sprünge bedingt sind, wird der Verlauf eines Phasenvergleichswerts vorteilhafterweise gesiebt und/oder geglättet. Das Ergebnis wird schließlich bevorzugt als Verlauf eines Phasenbeziehungswerts verwendet.

**[0070]** Bevorzugt wird der Verlauf eines Phasenbeziehungswerts als Steuerfunktion an den Zeitparameter eines Verzögerungsglieds geleitet.

**[0071]** Da es sich bei der Veränderung um eine nichtlineare Verarbeitung handelt, erfolgt diese zwecks Vermeidung von Aliasing-Artefakten bevorzugt unter mindestens vierfacher, insbesondere mindestens achtfacher Überabtastung der Samplingrate des Referenzsamples. Dazu wird das Referenzsample insbesondere mit künstlich erzeugten weiteren Werten, insbesondere Amplituden- und Phasenwerten, versehen. Zum Einsatz kommt dieses Verfahren insbesondere und häufig bei marktüblichen digitalen Emulationen von analogen Signalformern (etwa Distortion- oder Overdrive-Effekten) oder Dynamikprozessoren (etwa Kompressoren). Mit Vorteil wird nach der erfindungsgemäßen Anpassung jede der angepassten Audioquellen, insbesondere mittels Bandpassfilter, wieder auf die ursprüngliche Samplingrate oder Nyquistfrequenz bandbegrenzt und/oder mit der Samplingrate des verwendeten digitalen Mediums abgetastet.

**[0072]** Zur Veranschaulichung einer erfindungsgemäßen Lösung sei folgendes Beispiel aufgeführt: Das eingestrichene c einer Flöte sei nacheinander in den Dynamikstufen piano, mezzoforte und forte mit jeweils gleichzeitig drei Stereomikrofonpaaren in den drei Positionen nah, mittel und fern aufgenommen worden. Es liegen also insgesamt sechs Kanäle

vor. Als Referenzsample wird der linke Kanal des Mikrofonpaars "nah" der Dynamikstufe mezzoforte ausgewählt. Die beiden übrigen abhängigen Samples der Dynamikstufen piano und forte sollen nun an den Phasenverlauf der Mezzoforte-Stufe angepasst werden. Der jeweils entsprechende Kanal der Samples der Dynamikstufen forte bzw. piano wird nun mit dem Referenzsample verglichen. Es wird also der linke Kanal der Mikrofonposition "nah", mit dem linken Kanal "nah" des Referenzsamples korreliert, um jeweils den Verlaufs eines Phasenbeziehungswerts für jedes der abhängigen Samples zu erzeugen. Jeder der durch den Korrelator für jedes der abhängigen Sample erzeugte Verlauf eines Phasenbeziehungswerts steuert nun den Zeitparameter eines Verzögerungsgliedes, durch die das jeweilige abhängige Sample und zwar jeweils alle sechs Mikrofonkanäle in der Geschwindigkeit verändert werden. So werden die phasenkohärenten abhängigen Sample erzeugt.

**[0073]** Figur 1 zeigt schematisch und nicht beschränkend die Ermittlung eines Verlaufs eines Phasenbeziehungswerts ( $u$ ). In der Figur links werden das Referenzsampling ( $q$ ) und das abhängige Sample ( $s$ ) zunächst einem steifflankigen Bandpassfilter zur Isolation des Grundtons zugeführt. Um den Phasenversatz zwischen Referenzsample und abhängigen Sample zu ermitteln, werden die sinusförmigen Grundtonverläufe durch einen Komparator prozessiert, woraus eine Rechteckwelle resultiert, welche sodann differenziert wird. All dies geschieht nicht einmalig, sondern mehrfach, insbesondere kontinuierlich und/oder mindestens alle 100ms, insbesondere mindestens alle 10 ms, über eine beliebig lange Zeitdauer. Hierbei entsteht jeweils eine Folge von Diracimpulsen, welche sich gerade an den Nulldurchgängen der Samples befinden und je nach Richtung des Nulldurchgangs positiv bzw. negativ signiert sind. Der anschließende Komparator mit Gegenkopplung dient der Eliminierung der negativ signierten Impulse, welche fallenden Flanken entsprechen. Der Integrator im oberen Teil fungiert als Rampengenerator und wird durch die Impulse jeweils neu gestartet. Das an dessen Ausgang anliegende Signal gibt also die seit der letzten eingegangenen Flanke vergangene Zeit an. Dieses wird schließlich an den Eingang eines Abtast- und Halteglieders ( $s\&h$ ) gelegt, welches durch die Impulse aus dem unteren Schaltungsteil gesteuert wird. Am Ausgang liegt demnach schließlich stets (über die Dauer der Verfahrensführung) die zeitliche Differenz, den Phasenbeziehungswert ( $u$ ), zwischen den Flanken der beiden Eingangssignale der Gesamtschaltung an.

**[0074]** Figur 2 zeigt schematisch und nicht beschränkend einen Aufbau, der simultan alle Mikrofonssignale ( $s_0$  bis  $s_{N-1}$ ) einer einzigen Dynamikstufe, deren Phasenverlauf an das Referenzsignal ( $q$ ) anzugleichen ist, angleicht. Der Verlauf des Phasenbeziehungswertes ( $u$ ), dessen Ermittlung wie in Figur 1 dargestellt geschieht, wobei eins der Mikrofonssignale ( $s_0$  bis  $s_{N-1}$ ) als Signal ( $s$ ) verwendet wird, wird zunächst entklappt.

**[0075]** Anschließend wird das Signal tiefpassgefiltert und geclippt. Der Clipper, welcher Signalspitzen jenseits eines vorgegebenen Frequenzintervalls abschneidet, dient dem Zweck der Heraushebung von Spitzenartefakten. Die untere Intervallgrenze des Clippers ist identisch mit der Inversen Pufferzeit ( $T_{buf}$ ), die obere in diesem Beispiel fest auf  $10^8$ ms eingestellt.

**[0076]** Nach dieser kontinuierlichen Verarbeitung ergibt sich der Verlauf eines Phasenbeziehungswerts, auch Steuerungsfunktion ( $u$ ) (über die Dauer der Verfahrensführung).

**[0077]** Die einstellbare Pufferzeit ( $T_{buf}$ ) ist eine, insbesondere konstante, Grundverzögerung, insbesondere im Bereich von 5 bis 50 ms, für alle Mikrofonssignale ( $q$  und  $s_0$  bis  $s_{N-1}$ ), die zur Ermöglichung negativer Verzögerungszeiten benötigt wird und für die abhängigen Mikrofonssignale mit der Steuerungsfunktion addiert wird.

**[0078]** Die Summe aus Pufferzeit ( $T_{buf}$ ) und Verlauf des Phasenbeziehungswertes ( $u$ ) wird sodann einer Vielzahl Verzögerungsglieder ( $dly$ ) zugeführt, denen auch jeweils ein abhängiges Mikrofonssignale ( $s_0$  bis  $s_{N-1}$ ) zugeführt wird. Die Verzögerungsglieder verzögern die Ausgangssignale stets gegenüber den Eingangssignalen, wobei die Verzögerung nicht konstant ist. An den Ausgängen der Verzögerungsglieder ( $dly$ ) liegen dadurch phasenkohärente Signale ( $\bar{q}$ ,  $\bar{s}_0$  bis  $\bar{s}_{N-1}$ ) an.

**[0079]** Figur 3 zeigt schematisch und nicht beschränkend eine alternative latenzfreie Implementierung. Die in der ersten Implementierung der Figur 2 benötigte Pufferzeit führt zu einer ihr entsprechenden Latenzzeit der Gesamtschaltung. Für Anwendungen mit strikten Echtzeitanforderungen kann dies umgangen werden, indem anstelle von Verzögerungsgliedern Abspielmodule (hier metonymisch als "tape" bezeichnet) mit modulierbarer Abspielgeschwindigkeit, wie sie etwa in vielen Samplerumgebungen zur Verfügung stehen, verwendet werden.

**[0080]** Zur Beurteilung der durch die Erfindung möglichen Verbesserung wird im Nachfolgenden die folgende Metrik verwendet:

Gegeben seien nunmehr zwei Signale 0, 1 (Eingangsaudiosignale bzw. jeweils ein Abschnitt mit der gleichen Länge) mit identischer Periodendurchlaufszahl (Anzahl der Nulldurchgänge) sowie identischer mittlerer Periodendauer  $T$  (denn es handelt sich um Signale der gleichen Tonhöhe). Mit  $t^{(0)}$  und  $t^{(1)}$  gleich der Vektoren der Zeiten der Nulldurchgänge der Signale 0 und 1 ist mit einer geeigneten Norm ( $\|\cdot\|$ )

$$d_{\parallel \cdot \parallel} (t^{(0)}, t^{(1)}) := \left\| \frac{t^{(1)} - t^{(0)}}{T} \right\|$$

ein Maß für die Phaseninkohärenz der Signale 0, 1.

**[0081]** Hier wird nun die RMS Norm gewählt, wobei RMS zwecks Dimensionsinvarianz durch die Anzahl der Periodendurchläufe normalisiert sind.

**[0082]** Entsprechend der in der Signalverarbeitung üblichen Vorgehensweise ist es ratsam, mit einer gefensterten RMS-Norm zu arbeiten, was mathematisch als eine Kombination von RMS und PEAK gedeutet werden kann. Als Fenstergröße wurde ein RMS (root mean square)-Fenster entsprechend der in den Figuren 4 bis 11 auf der horizontalen Achse in ms dargestellten Zeitdauern gewählt. Die Reihenfolge der Zeilen entspricht der Reihenfolge der Figuren.

**[0083]** Als Maß für die Güte des Phasenangleichungsverfahrens bietet sich nun das Verhältnis zwischen gemessener Phasenkohärenz vor und nach der Verarbeitung durch das Verfahren oder die relative Veränderung der Phasenkohärenz an.

**[0084]** Es wurden nun verschiedene Schallereignisse analysiert. Dazu wurde jeweils eine multimikrofonierte Schallquelle an drei unterschiedlichen Stereomikrofonpositionen sowie an einer weiteren einkanaligen Referenzmikrofonposition aufgezeichnet. Dadurch wurden drei Stereo und eine Monospur erhalten. Exemplarisch wurde der linke Kanal eines Stereomikrofonpaares (Signal 1) als zu bearbeitendes Eingangsaudiosignal und das Signal des Referenzmikrofons als Referenzaudiosignal (Signal 0) ausgewählt und die Phasenkohärenz im Sinne der besprochenen Metrik zwischen Referenz- (Signal 0) und Eingangsaudiosignal (Signal 1) bzw. bearbeitetem Eingangsaudiosignal (Signal 1') bestimmt. Zudem wurde zum Vergleich mit dem Stand der Technik eine Bearbeitung des Eingangsaudiosignals (Signal 1) mittels konstanter Zeitverschiebung nach der JP 2001100756 A vorgenommen (Signal 1").

	d(0,1)	d(0,1')	d(0,1') / d(0,1)	(d(0,1)-d(0,1')) / d(0,1)	d(0,1")	d(0,1") / d(0,1)	(d(0,1)-d(0,1")) / d(0,1)
CE   ff   45 (a1)	0,2173	0,0608	28%	72%	0,4	184%	-84%
VL   f   55 (g2)	0,4215	0,0545	13%	87%	0,235	56%	44%
VL   f   60 (c3)	0,3983	0,1926	48%	52%	0,264	66%	34%
VL   f   69 (a3)	0,2737	0,0712	26%	74%	0,2879	105%	-5%
VL   f   72 (c4)	0,2524	0,1387	55%	45%	0,3141	124%	-24%
VL   f   77 (f4)	0,2953	0,1952	66%	34%	0,2721	92%	8%
VL   f   78 (f#4)	0,2993	0,143	48%	52%	0,2502	84%	16%
VL   f   84 (c5)	0,2465	0,1183	48%	52%	0,3471	141%	-41%

**[0085]** Dabei wurde in Cello (CE) und eine Violine (VL) als Schallquelle verwendet. Es wurden verschiedene Dynamikstufen (f und ff) bei verschiedenen Tonhöhen (Angabe in Midi-Tonhöhenstufe und in Klammer musikalischer Notation) getestet.

**[0086]** Zu erkennen ist, dass bei statischer Korrektur teilweise eine Verbesserung und teilweise eine Verschlechterung erreicht wird, die stark schwankt. Mit dem erfindungsgemäßen Verfahren wird immer eine Verbesserung erreicht, die die Verbesserung gegenüber dem Stand der Technik mindestens um 18%-Punkte übersteigt.

**[0087]** Der Toningenieur hat für die Verarbeitung die oben beschriebenen Parameter "Glättungsgrad" und "Vorauschau" wie oben erläutert nach Gehör in Hinblick auf ein subjektiv optimales Verhältnis zwischen hörbarer Phasenkohärenz (d.h., Abwesenheit von Kammfiltereffekten) einerseits und Artefaktfreiheit andererseits eingestellt, ohne dabei eine Metrik zu beachten.

**[0088]** Die Figuren 4 bis 11 geben die Verläufe der RMS Werte (RMS Werte auf der vertikalen Achse) mit einem Fenster von jeweils 500 Periodendurchläufen der Messungen, die der vorstehenden Tabelle zu Grunde liegen, der Reihenfolge der Zeilen der Tabelle nach wieder. Auf der horizontalen Achse ist die Zeit in ms abgetragen. Zu erkennen ist hier, dass über weite Teile der Signaldauer eine Verbesserung erreicht werden kann. Dabei ist das Signal A das Signal 1, das Signal B das Signal 1" und das Signal C das Signal 1'.

## Patentansprüche

1. Verfahren zur Erstellung einer Gruppe phasenkohärenter Audioquellen ( $\bar{q}$ ,  $\bar{s}_0$  bis  $\bar{s}_{N-1}$ ), insbesondere Audiosignale oder Samples, umfassend:
  - a) Bereitstellen mindestens jeweils einer Eingangsaudioquelle ( $q$ , so bis  $s_{N-1}$ ), insbesondere Audiosignale oder Samples, eines ersten und eines zweiten Tonerereignisse, der gleichen Tonhöhe eines einzigen realen Musikinstruments,
  - b) Auswählen einer Audioquelle, insbesondere einer der Eingangsaudioquellen des ersten Tonerereignisses, als Referenzsample ( $q$ ),
  - c) Vergleich des Phasenverlaufs des Referenzsamples ( $q$ ) mit dem Phasenverlauf mindestens einer weiteren der Eingangsaudioquellen ( $s_0$  bis  $s_{N-1}$ ) des zweiten Tonerereignisses über eine erste Audioquellendauer des Referenzsamples zur Erzeugung einer Phasenbeziehungsinformation ( $u$ ),
  - d) Verändern der Geschwindigkeit und/oder Zeitbasis der mindestens einen weiteren Eingangsaudioquelle ( $s_0$  bis  $s_{N-1}$ ) über eine zweite Dauer der mindestens einen weiteren Eingangsaudioquelle auf Basis der Phasenbeziehungsinformation ( $u$ ), **dadurch gekennzeichnet, dass** die Phasenbeziehungsinformation ein variierender Verlaufes eines Phasenbeziehungswertes ist und dass die Geschwindigkeit und/oder Zeitbasis über die zweite Dauer variierend abhängig vom Verlauf des Phasenbeziehungswertes verändert wird.
2. Verfahren nach Anspruch 1, wobei die erste Dauer der zweiten Dauer entspricht und/oder die erste Dauer länger ist als die zweite Dauer.
3. Verfahren nach einem der vorstehenden Ansprüche, wobei der Wert des Verlaufs des Phasenbeziehungswertes an einem Zeitpunkt des Referenzsamples zumindest teilweise den Phasenverlauf des Referenzsamples und/oder den Phasenverlauf der einer weiteren der Audioquellen, vor, an und/oder nach dem gleichen Zeitpunkt im Referenzsample berücksichtigt, wobei die Zeitpunkte insbesondere relativ zum Start des jeweiligen Tonerereignisses gemessen sind.
4. Verfahren nach einem der vorstehenden Ansprüche, wobei das Verfahren für mehrere zweite Tonerereignisse, insbesondere gleichzeitig, mit einem gemeinsamen ersten Tonerereignis und einem gemeinsamen Referenzsample durchgeführt wird und für jedes zweite Tonerereignis eine Phasenbeziehungsinformation erzeugt wird und jede Phasenbeziehungsinformation zur Veränderung der Geschwindigkeit mindestens eines insbesondere mehrerer, Eingangsaudioquelle des jeweiligen zweiten Tonerereignisses verwendet wird.
5. Verfahren nach einem der vorstehenden Ansprüche, wobei die Schritte c) und d) mittels mindestens einem, insbesondere mittels mindestens genau einem, gegengekoppeltem Regelkreis ausgeführt werden, wobei insbesondere eine Messschleife pro zweitem Tonerereignis zur Ermittlung der Phasenbeziehungsinformation verwendet wird und/oder eine Steuerungsschleife pro weitere Eingangsaudioquelle zur Veränderung der Geschwindigkeit verwendet wird.
6. Verfahren zum zeitgleichen Abspielen mindestens jeweils einer Audioquelle, insbesondere Audiosignal oder Sample, zweier Tonerereignisse, insbesondere von mindestens 500ms Länge, der gleichen Tonhöhe des gleichen realen Instruments, wobei die Audioquellen jeweils, insbesondere über mindestens 500ms, zumindest teilweise gleichzeitig, insbesondere über mindestens 500ms, phasenkohärent abgespielt werden und insbesondere nach Anspruch 1 erzeugt werden und/oder erzeugt sind.
7. Virtuelles Instrument aufweisend und/oder eingerichtet zur Erzeugung, insbesondere nach Anspruch 1, mindestens einer Gruppe phasenkohärenter Samples, umfassend mindestens zwei Samples jeweils eines Tonerereignisse, insbesondere von mindestens 500ms Länge, der gleichen Tonhöhe eines, insbesondere des gleichen, realen Musikinstruments,

wobei das virtuelle Instrument eingerichtet ist, mindestens zwei der phasenkohärenter Samples gleicher Tonhöhe zumindest teilweise gleichzeitig, insbesondere über mindestens 500ms, phasenkohärent abzuspielen, wobei die Samples eine Länge, insbesondere von mindestens 500ms, aufweisen, in denen sie sich nicht wiederholen, **dadurch gekennzeichnet, dass** mindestens ein Sample der Gruppe eine über dessen Zeitverlauf variierende Geschwindigkeitsänderung aufweist und/oder mindestens ein Sample der Gruppe ein bearbeitetes Sampling ist, das ein über seine zeitliche Länge, insbesondere ohne Ein- und Ausschwingzeiten, bestimmtes RMS über die auf die mittlere Periode im Intervall normalisierten Abweichungen von der mittleren Periode des Intervalls und/oder von der Periode eines Referenzsamples aus der Gruppe von Samples von weniger als 75% des ebenso bestimmten RMS des unbearbeiteten Sample, aus dem das bearbeitete Sampling hervorgegangen ist und/oder des mittels eines statischen Zeitversatzes korrigierten Samples, aus dem das bearbeitete Sampling hervorgegangen ist, aufweist.

8. Gruppe von Samples, wobei mindestens eines der Sample phasenkorrigiert ist und wobei die Samples solche unterschiedlicher Dynamiken einer Tonhöhe eines realen Musikinstruments darstellen, wobei die Samples jeweils eine zeitliche Länge, insbesondere von mindestens 500ms, aufweisen, wobei sich die Samples über die zeitliche Länge nicht wiederholen und wobei mindestens ein Sample der Gruppe eine über dessen Zeitverlauf variierende Geschwindigkeitsänderung aufweist und/oder mindestens ein Sample der Gruppe ein bearbeitetes Sampling ist, das ein über seine zeitliche Länge, insbesondere ohne Ein- und Ausschwingzeiten, bestimmtes RMS über die auf die mittlere Periode im Intervall normalisierten Abweichungen von der mittleren Periode des Intervalls und/oder von der Periode eines Referenzsamples aus der Gruppe von Samples von weniger als 75% des ebenso bestimmten RMS des unbearbeiteten Samples, aus dem das bearbeitete Sampling hervorgegangen ist und/oder des mittels eines statischen Zeitversatzes korrigierten Samples, aus dem das bearbeitete Sampling hervorgegangen ist, aufweist und/oder die Gruppe Informationen zum Phasenverlauf der Samples über die Länge, insbesondere von mindestens 500ms, aufweist.
9. Verwendung einer Gruppe von Samples, insbesondere nach Anspruch 5, wobei mindestens ein Sampling der Gruppe phasenkorrigiert ist und/oder mindestens ein Sampling der Gruppe Informationen zum Phasenverlauf aufweist, wobei die Samples mindestens zweier unterschiedlicher Dynamiken einer Tonhöhe eines realen Musikinstruments darstellen, wobei die Samples eine Länge, insbesondere von mindestens 500ms, aufweisen, in denen mindestens eins phasenkorrigiert ist und/oder Informationen zum Phasenverlauf aufweist und wobei sich die Samples über die Länge nicht wiederholen, zur phasenkorrigierten und zumindest teilweise gleichzeitigen Wiedergabe der Samples einer Tonhöhe mit mindestens zwei unterschiedlichen Dynamiken.

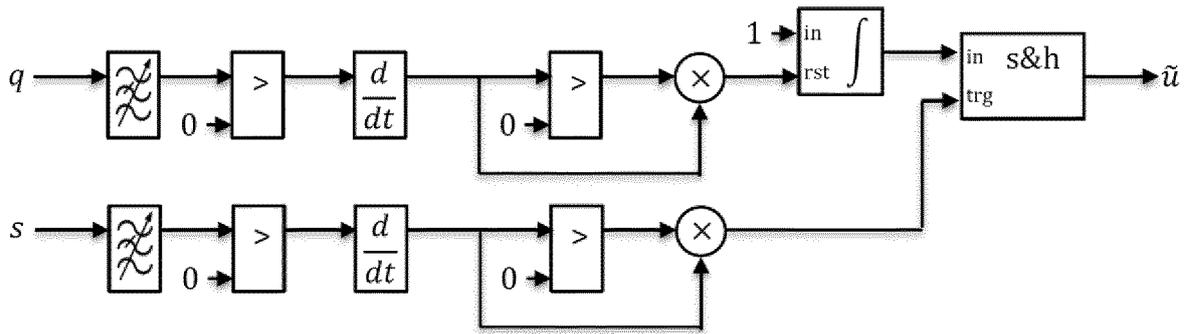


Fig. 1

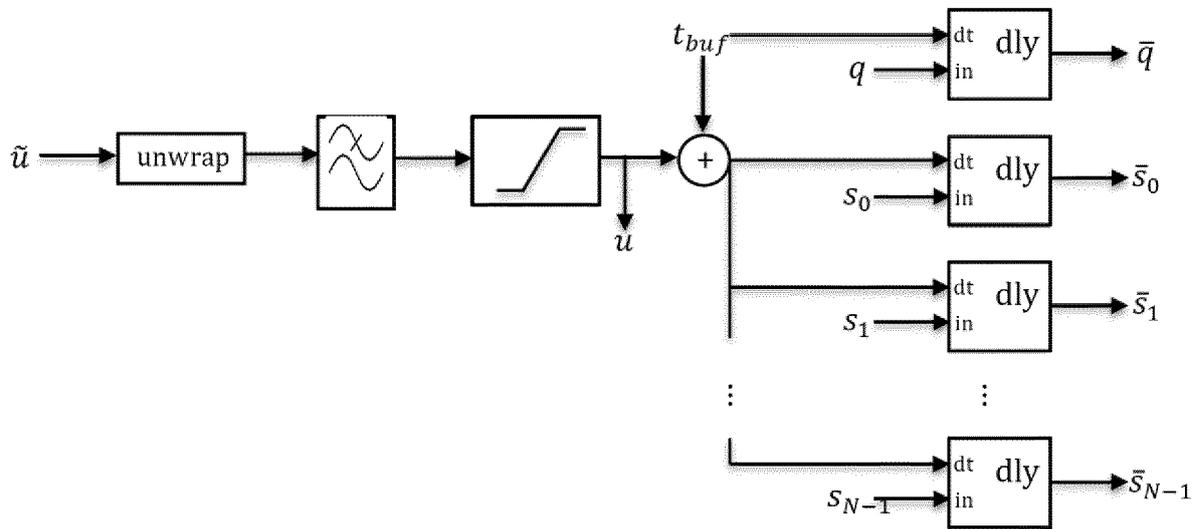


Fig. 2

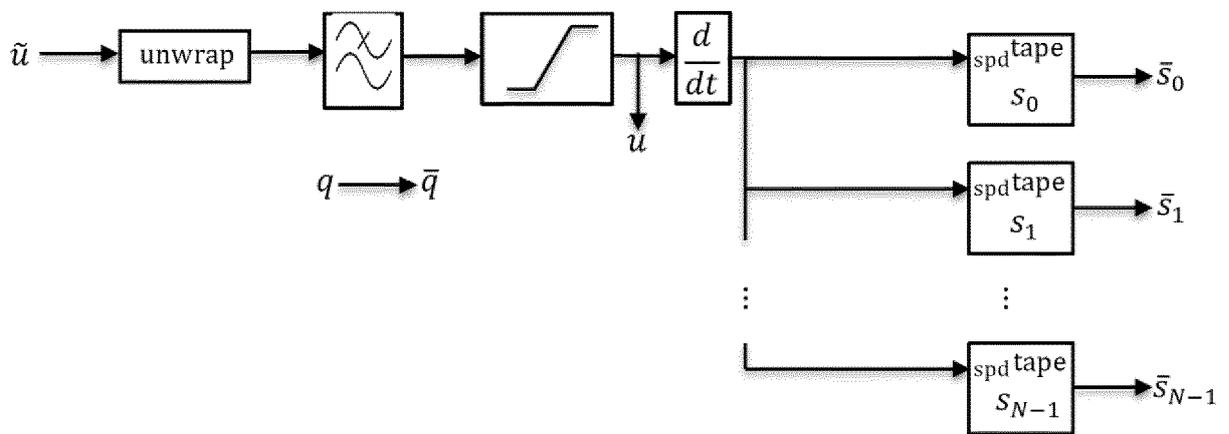


Fig. 3

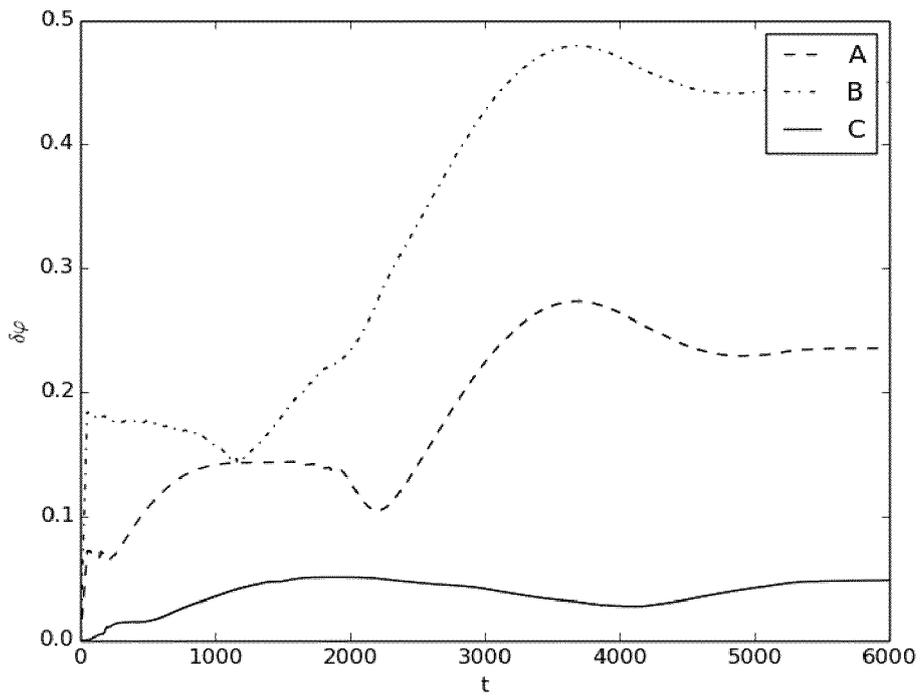


Fig. 4

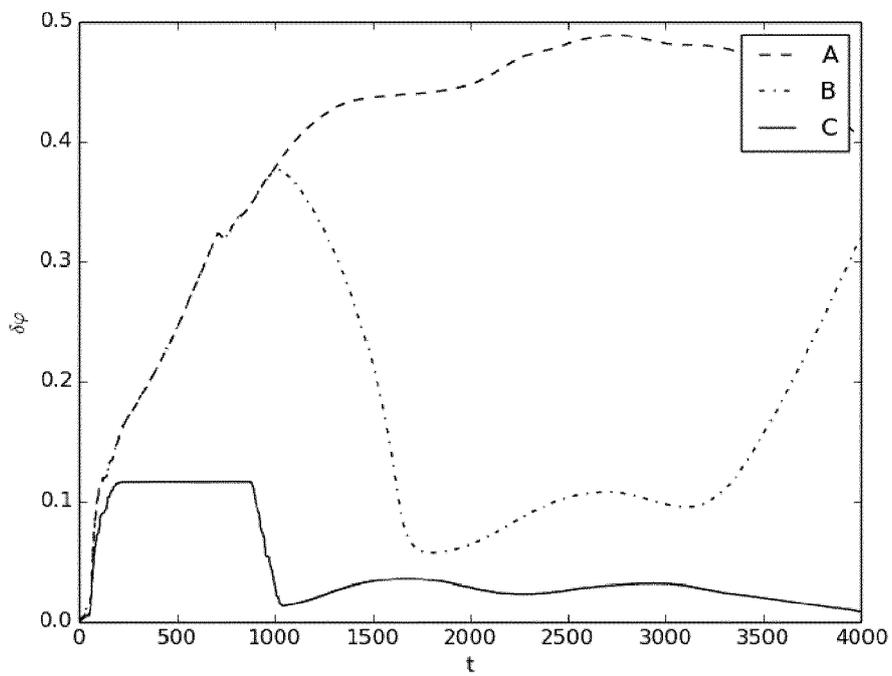


Fig. 5

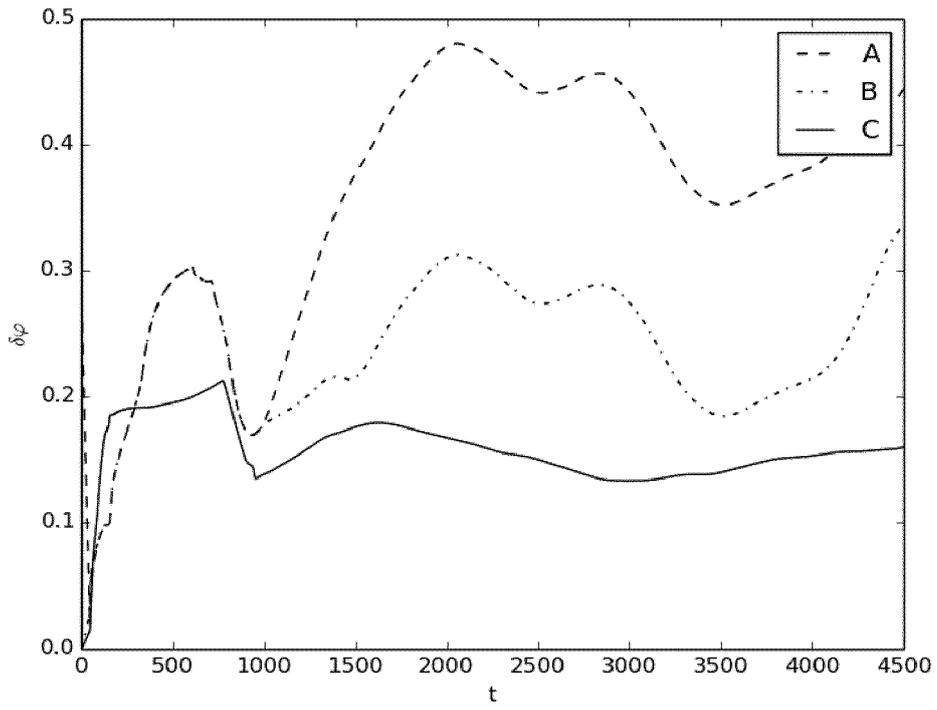


Fig. 6

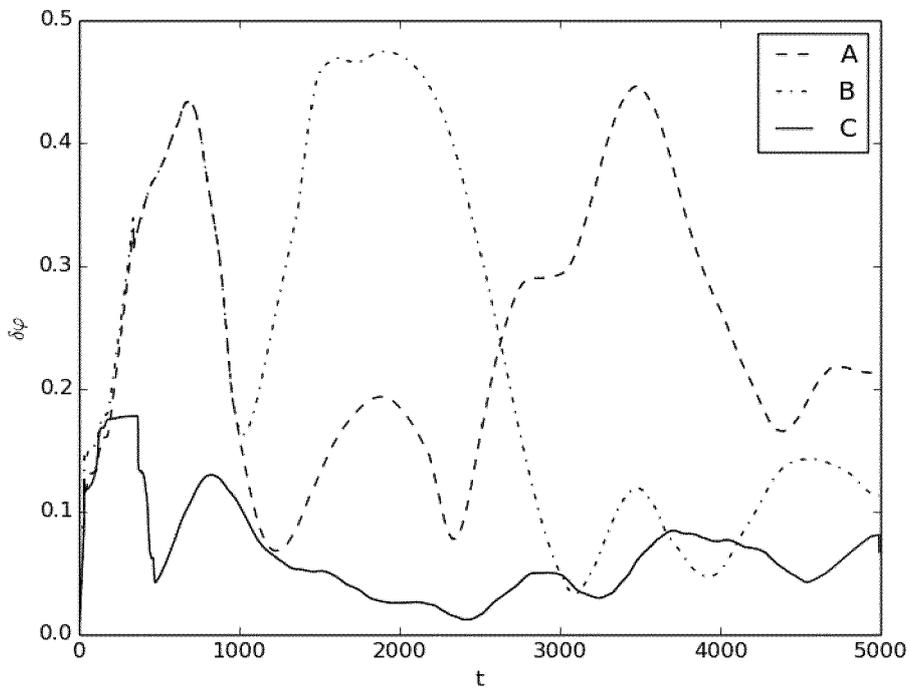


Fig. 7

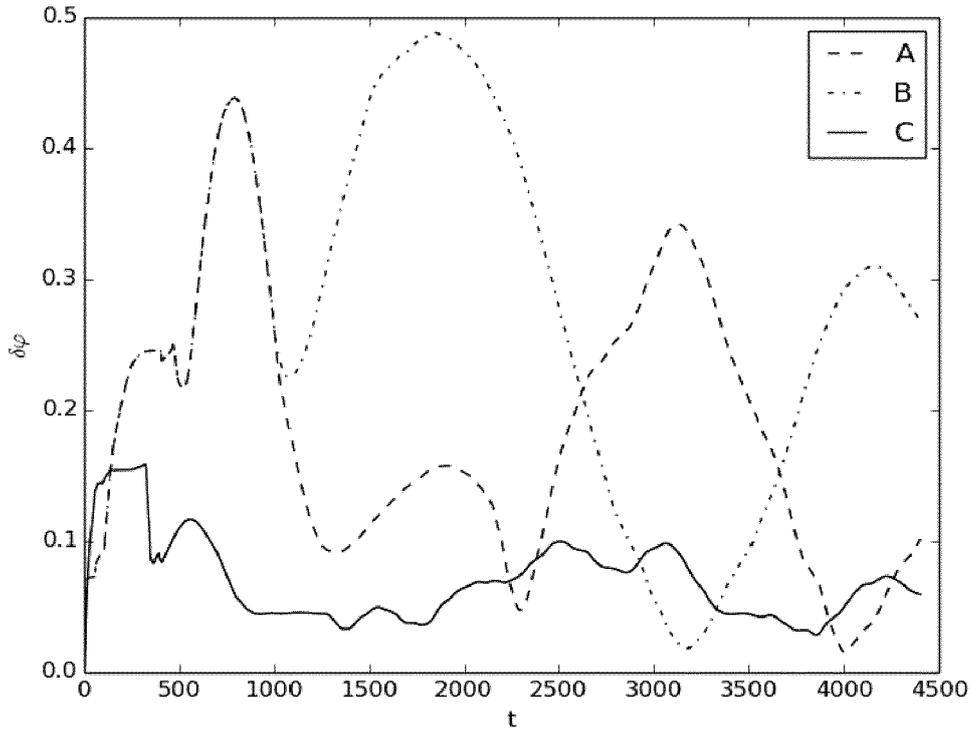


Fig. 8

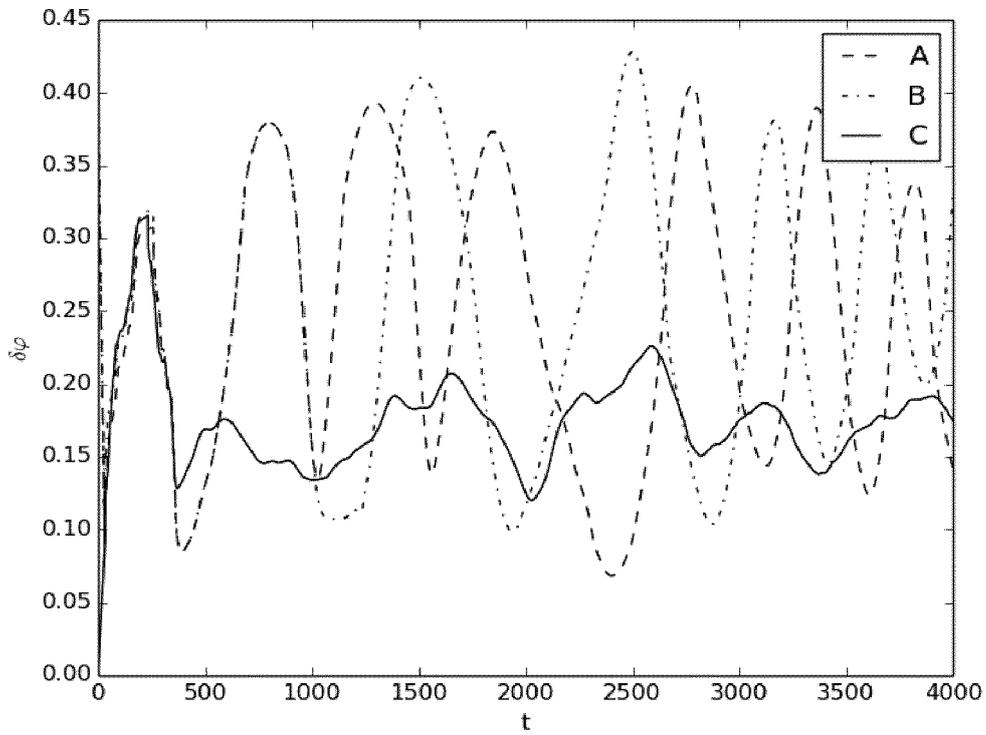


Fig. 9

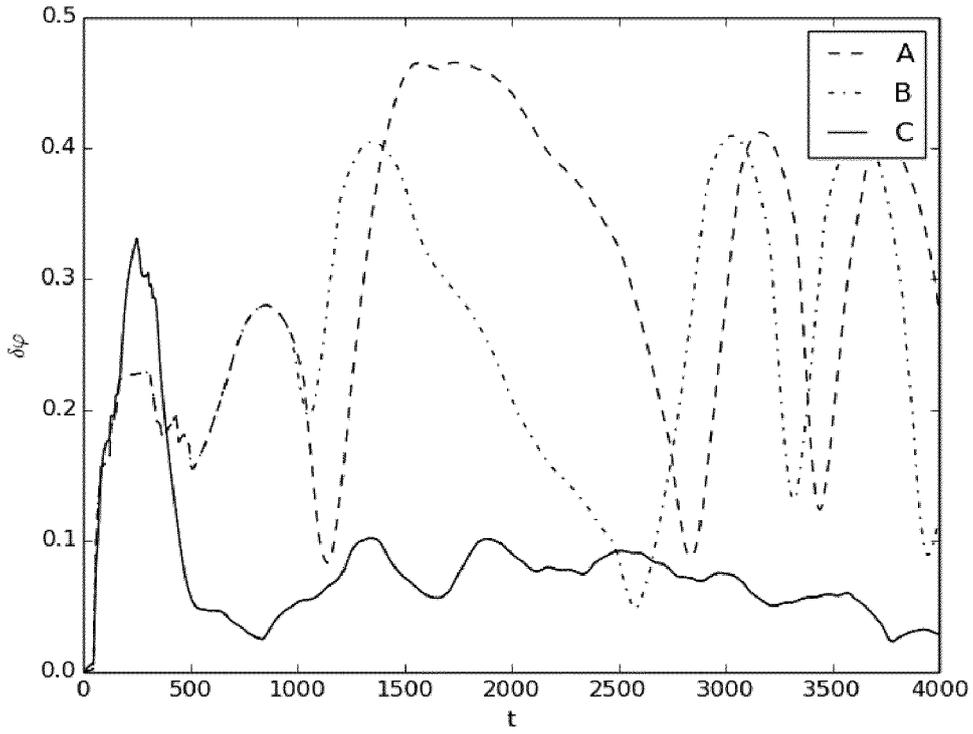


Fig. 10

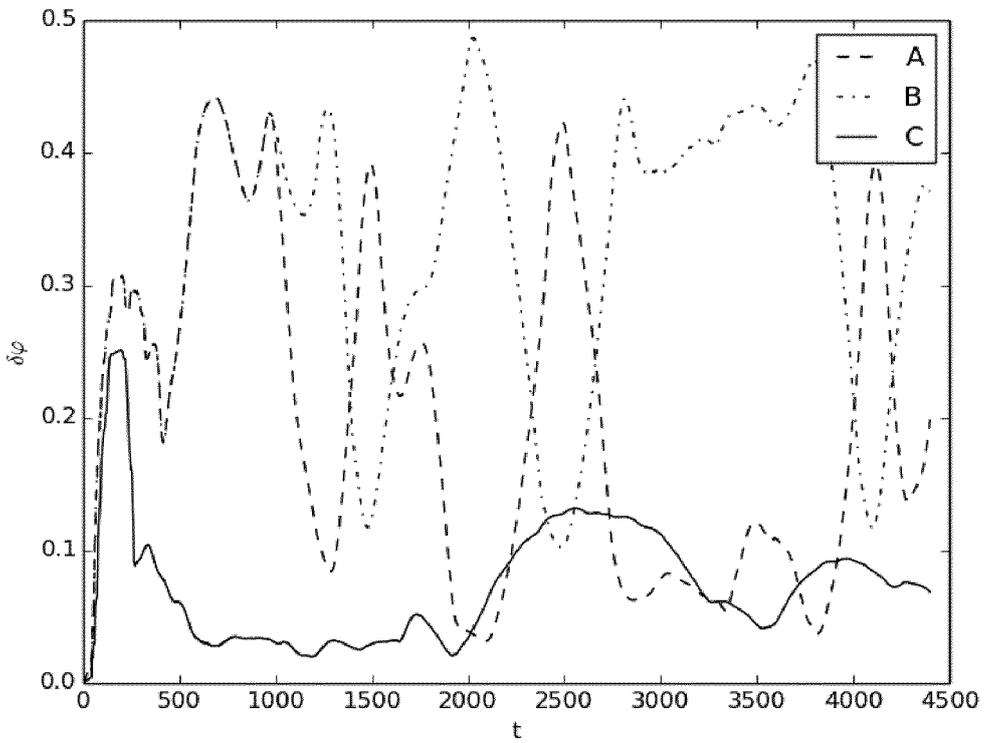


Fig. 11



EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung

EP 22 15 3060

5

10

15

20

25

30

35

40

45

50

55

EINSCHLÄGIGE DOKUMENTE			
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betrifft Anspruch	KLASSIFIKATION DER ANMELDUNG (IPC)
X	US 4 779 505 A (SUZUKI HIDEO [JP]) 25. Oktober 1988 (1988-10-25)	1-6	INV. G10H7/00
A	* Spalte 2, Zeile 62 - Spalte 9, Zeile 43 * * Abbildungen 6, 7 *	7-9	G10H7/04
A	US 2012/243711 A1 (FUJITA HIROAKI [JP] ET AL) 27. September 2012 (2012-09-27) * Absätze [0002], [0008] - [0046]; Abbildungen 3, 9, 10, 13, 14 *	1-9	
A	US 2017/372711 A1 (SHIRAHAMA TARO [JP]) 28. Dezember 2017 (2017-12-28) * Absätze [0002] - [0005], [0120] - [0127]; Abbildungen 11-15 *	1-9	
A	DE 40 08 872 A1 (WERSI GMBH & CO [DE]) 26. September 1991 (1991-09-26) * Spalte 8, Zeile 24 bis Spalte 10, Zeile 41 *	1-9	
			RECHERCHIERTE SACHGEBIETE (IPC)
			G10H
Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt			
Recherchenort <b>München</b>		Abschlußdatum der Recherche <b>24. Juni 2022</b>	Prüfer <b>Müller, Achim</b>
KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur		T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument ..... & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument	

1  
EPO FORM 1503 03.82 (P04C03)

**ANHANG ZUM EUROPÄISCHEN RECHERCHENBERICHT  
 ÜBER DIE EUROPÄISCHE PATENTANMELDUNG NR.**

EP 22 15 3060

5 In diesem Anhang sind die Mitglieder der Patentfamilien der im obengenannten europäischen Recherchenbericht angeführten Patentedokumente angegeben.  
 Die Angaben über die Familienmitglieder entsprechen dem Stand der Datei des Europäischen Patentamts am  
 Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

24-06-2022

Im Recherchenbericht angeführtes Patentedokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
<b>US 4779505 A</b>	<b>25-10-1988</b>	<b>KEINE</b>	
-----			
<b>US 2012243711 A1</b>	<b>27-09-2012</b>	<b>CN 102740216 A</b>	<b>17-10-2012</b>
		<b>EP 2503791 A2</b>	<b>26-09-2012</b>
		<b>JP 6102063 B2</b>	<b>29-03-2017</b>
		<b>JP 2012217146 A</b>	<b>08-11-2012</b>
		<b>US 2012243711 A1</b>	<b>27-09-2012</b>
-----			
<b>US 2017372711 A1</b>	<b>28-12-2017</b>	<b>CN 104867487 A</b>	<b>26-08-2015</b>
		<b>EP 2911147 A1</b>	<b>26-08-2015</b>
		<b>EP 3261086 A1</b>	<b>27-12-2017</b>
		<b>JP 6090204 B2</b>	<b>08-03-2017</b>
		<b>JP 2015158527 A</b>	<b>03-09-2015</b>
		<b>US 2015243291 A1</b>	<b>27-08-2015</b>
		<b>US 2017372711 A1</b>	<b>28-12-2017</b>
-----			
<b>DE 4008872 A1</b>	<b>26-09-1991</b>	<b>KEINE</b>	
-----			

EPO FORM P0461

Für nähere Einzelheiten zu diesem Anhang : siehe Amtsblatt des Europäischen Patentamts, Nr.12/82

**IN DER BESCHREIBUNG AUFGEFÜHRTE DOKUMENTE**

*Diese Liste der vom Anmelder aufgeführten Dokumente wurde ausschließlich zur Information des Lesers aufgenommen und ist nicht Bestandteil des europäischen Patentdokumentes. Sie wurde mit größter Sorgfalt zusammengestellt; das EPA übernimmt jedoch keinerlei Haftung für etwaige Fehler oder Auslassungen.*

**In der Beschreibung aufgeführte Patentdokumente**

- US 9805702 B1 [0002]
- US 20120243711 A1 [0003]
- US 20170372711 A1 [0004]
- JP 2001100756 A [0005] [0084]
- WO 2018055892 A1 [0006]
- DE 4008872 A1 [0007]