

Kapitel 13

Audiobearbeitung

Hans-Joachim Maempel, Stefan Weinzierl und Peter Kaminski

13.1	Editierung und Montage	721
13.1.1	Historische Entwicklung	721
13.1.2	Praxis und Funktion	723
13.2	Klanggestaltung	725
13.2.1	Fader und Mute	725
13.2.2	Panoramaregler	726
13.2.3	Stereo-Matrix	729
13.2.4	Kompressor	730
13.2.5	Expander	739
13.2.6	Kombinierte Regelverstärker	743
13.2.7	Filter	744
13.2.8	Verzerrer	747
13.2.9	Enhancer	748
13.2.10	Delayeffekte	748
13.2.11	Nachhall	751
13.2.12	Pitch Shifting	755
13.2.13	Time Stretching	756
13.2.14	Phaser	756
13.2.15	Ringmodulator	757
13.2.16	Tremolo und Vibrato	758
13.2.17	Leslie-Kabinett (rotary speaker)	758
13.2.18	Wah-Wah	759
13.2.19	Vocoder	760
13.2.20	Mehrfach-Bearbeitung und komplexe Algorithmen	762
13.3	Klangrestauration	763
13.3.1	Bearbeitungswerkzeuge	763
13.3.2	Klassifizierung von Störungen	764
13.3.3	Rauschen	764
13.3.4	Scratches, Clicks und Crackles	766
13.3.5	Frequenzgangsbeschneidung	768
13.3.6	Pop-Geräusche	769
13.3.7	Übersteuerungen und Verzerrungen	769
13.3.8	Netz-Störungen	770

13.3.9	Azimuthfehler	771
13.3.10	Drop-Outs	772
13.3.11	Digitale Störungen	772
13.3.12	Störgeräusche	773
13.3.13	Praxis	774
13.4	Funktionen und ästhetische Ziele	775
13.4.1	Physik und Psychologie	776
13.4.2	Gestalten und Interpretation	778
13.4.3	Klangästhetische Maximen	779
13.4.4	Aktivierung	780
13.4.5	Assoziation	780
13.4.6	Robustheit und Lautheit	781
	Literatur	781

Audioinhalte werden bei der Aufnahme, Übertragung und Wiedergabe in der Regel auch bearbeitet, da sowohl künstlich erzeugte als auch in der Natur aufgenommene Klänge zunächst oft nur unzureichend für die mediale Verbreitung geeignet sind. Im wesentlichen gibt es dabei vier Zielsetzungen, die je nach Inhalten und Medien unterschiedlich stark gewichtet sind und ineinander greifen:

- Die Produktion einer vorlagengetreuen und idealisierten akustischen Darbietung oder Szene (z. B. die Eliminierung von Störungen oder Spielfehlern)
- Die Interpretation der aufgenommenen Inhalte durch technische Mittel (z. B. die Gewichtung oder Effektbearbeitung von Musikinstrumenten)
- Die Anpassung an die technischen Erfordernisse der Übertragungskette (z. B. korrekte Aussteuerung)
- Die Anpassung an die Rezeptionssituationen, -gewohnheiten und -erwartungen der Zielgruppe (z. B. die Einhaltung einer sinnvollen Programmdynamik)

Entsprechend umfangreich wird Audiotbearbeitung heute eingesetzt: in allen elektronischen Medien, während nahezu aller Produktionsschritte und bei allen Audioinhalten. Bis in die 1970er Jahre hinein funktionierten Audiotbearbeitungsmittel ausschließlich auf der Basis analoger Signalverarbeitung. Als Ergebnis der technologischen Entwicklung überwiegt heute die digitale Realisation, entweder als Stand-Alone-Gerät (Hardware), als Software-Anwendung, als sogenanntes Plug-in oder als Hardware-Plug-in-Kombination. Plug-ins sind spezialisierte Software-Komponenten, die über standardisierte Schnittstellen (z. B. AU, VST) in integrierte Audioproduktionsumgebungen (digital audioworkstations, DAWs) eingebunden werden können. Sie bestehen aus einem Algorithmus zur Berechnung der Audiodaten und einer grafischen Benutzeroberfläche. Einfache Bearbeitungsmittel wie Equalizer und Regelverstärker sowie zunehmend auch Effekte wie Delays oder Nachhall sind häufig in die Kanalzüge professioneller Mischpulte integriert. Durch digitale Signalverarbeitung wurden bestehende Audiotbearbeitungsmittel auf höherem qualitativem Niveau realisierbar (insbesondere künstlicher Nachhall) sowie neue Effekte ermöglicht, z. B. Pitch Shifting und Time Stretching mit Formantkorrektur. Gegenüber tradi-

tionellen Verfahren der Signalbearbeitung durch Filter mit signalunabhängig festgelegten Parametern hat die digitale Audiosignalverarbeitung in jüngerer Zeit neue Verfahren erschlossen durch den Einsatz von Faltungsoperationen in Echtzeit und anderer Frequenzbereichsverfahren, sowie durch die dynamische Parametrisierung von Bearbeitungsoperationen auf der Basis von Merkmalsanalysen (feature extraction).

13.1 Editierung und Montage

Professionelle Audioproduktionen (Soundtracks zum Film, Hörspiele, Musikaufnahmen) sind in der Regel das Ergebnis einer Montage von einzelnen Aufnahmen, deren Einspielung sich nicht notwendigerweise an der zeitlichen Struktur des Endprodukts orientiert. Im Bereich der Populärmusik (außer Jazz) wird überwiegend im sog. Overdubbing-Verfahren produziert. Einzelne Stimmen eines Arrangements werden nacheinander auf unterschiedliche Spuren aufgenommen, Korrekturen erfolgen durch destruktives oder nichtdestruktives Überschreiben von kurzen Sequenzen auf dem mehrspurigen Audiomaterial (punch in – punch out). Im Bereich der klassischen Musik ebenso wie bei Sprachaufnahmen werden dagegen, ähnlich wie beim Film, zunächst mehrere Versionen (Takes) eines Stücks oder einzelner Passagen eines Stücks aufgenommen. In einem anschließenden Editierungsprozess wird das aufgenommene Tonmaterial zu einem organischen Ablauf montiert. So entsteht die Illusion einer musikalischen Aufführung, die in dieser Form nie stattgefunden hat und vielleicht nie stattfinden könnte.

13.1.1 Historische Entwicklung

Während der Schnitt beim Film, der ja nichts anderes ist als eine Aneinanderreihung von Fotografien, keine speziellen Anforderungen an den Bildträger stellt und daher schon immer Teil des Produktionsprozesses war, waren beim Ton zunächst nur bestimmte Aufzeichnungsmedien geeignet, eine Aufzeichnung neu zu montieren und dabei unhörbare Übergänge zu schaffen. Während Nadeltonträger (Zylinder, Wachs- und Schellackplatte) für eine solche Montage ungeeignet sind, wurde bereits Ende der 1920er Jahre das neu entwickelte Lichtton-Verfahren erstmals für Hörspiel-Montagen eingesetzt. Für *Weekend*, eine akustische Montage von Alltagsgeräuschen einer Großstadt am Wochenende, übertrug der Komponist Walter Ruttmann, der auch die Filmmusik zu *Berlin – Sinfonie einer Großstadt* (1927) komponiert hatte, das Collage-Prinzip auf den Ton. 2000 m Filmtón wurden am Schneidetisch auf 250 m gekürzt, um ein 11'20" langes Stück aus 240 Einzelsegmenten zu erhalten, das 1930 in der Sendung *Hörspiele auf Tonfilm* zum ersten Mal im Rundfunk gesendet wurde (Vowinckel 1995).

Das Magnetband wurde in Deutschland seit Einführung der HF-Vormagnetisierung im Jahr 1940 für Musikproduktionen, Mitschnitte und für den Sendebetrieb

eingesetzt. Außerhalb Deutschlands hielt es erst nach 1945 Einzug. In den USA wurden seit 1948 beim Rundfunksender ABC und beim Decca-Label Magnetbandgeräte eingesetzt (Gooch 1999), seit 1949 wurde in England bei Decca (Culshaw 1981), ab 1950 auch bei E.M.I. auf Magnetband aufgezeichnet (Martland 1997). Das Magnetband konnte und musste, da es zunächst sehr rissanfällig war, geklebt werden. Dabei wurde zunächst „nass“ geklebt, d.h. bei beiden mit einer unmagnetischen Schere geschnittenen Enden des Bandes wurde die Trägerschicht durch ein Lösungsmittel leicht angelöst, das Band einige mm überlappt und durch anpressen verklebt. Seit Anfang der 1950er Jahre waren Klebelehren auf dem Markt, mit denen der Schnittwinkel der Bandenden exakt eingestellt werden konnte, sowie Klebestreifen zum Hinterkleben des Bandmaterials. Bei Interpreten stieß die von den Musikproduzenten in den 1960er Jahren bereits intensiv genutzte Möglichkeit der Montage, wie das immer größere Repertoire an Audiotbearbeitungsverfahren im Allgemeinen, auf ambivalente Reaktionen. Während Interpreten wie der kanadische Pianist Glenn Gould (1932–1982) die Montage als Erweiterung der interpretatorischen Gestaltungsmöglichkeiten verstand und sich zugunsten der Studioproduktion zunehmend aus dem Konzerleben zurückzog (Gould 1966), erlaubte der rumänische Dirigent Sergiu Celebidache (1912–1996) nur noch Live-Übertragungen seiner Konzerte, wodurch die zeitliche Struktur unangetastet blieb (Celebidache 1985).

Schnitt- und Klebelehren für analoges Magnetband erlaubten einen wenige Millimeter langen Schrägschnitt, entsprechend einer Überblenddauer (crossfade) von etwa 10 ms bei einer Bandgeschwindigkeit von 38 cm/s. Vereinzelt arbeitete man in den 1970er Jahren mit bis zu 20 cm langen Schnitten, entsprechend einer Überblendzeit von über 500 ms, wobei das Band häufig in „Schwalbenschwanz“-Form geschnitten wurde, um die beim Schrägschnitt aufgrund der Spurlage auftretenden, bei kurzen Schnitten aber in der Regel unhörbaren Störungen des Stereo-Panoramas zu vermeiden. Bei Mehrspuraufnahmen wäre der Zeitversatz zwischen den Spuren bei solch flachen Schnittwinkeln allerdings zu groß gewesen, weshalb man bereits bei 8-Spuraufnahmen mit „Zickzackbrettern“ arbeitete, die parallele Schrägschnitte in das Band stanzen.

Bereits das erste im kommerziellen Einsatz befindliche digitale Aufzeichnungssystem (Soundstream, 1976), das mit Festplatten und Bandsystemen als Aufzeichnungsmedium arbeitete, übertrug die Editiermöglichkeiten des analogen Tonbandes auf die Digitaltechnik, wodurch Funktionen wie Wellenformdarstellung, Crossfades und Cut, Copy und Paste eingeführt wurden. Zunächst setzten sich in der Musikproduktion bandgestützte Systeme durch, die mit einem Umkopierschnitt arbeiteten. Hier wurden zwei Zuspeler so zueinander synchronisiert, dass an einem definierten Schnittpunkt von einem zum anderen übergeblendet und das Ergebnis auf einem dritten Gerät im Aufnahmemodus aufgezeichnet werden konnte. Ein kleiner Bereich um die Schnittstelle konnte durch Editoren wie das 1980 eingeführte Modell DAE 1100 der Fa. Sony im RAM vorgehört werden. Als Tonträger konnten verschiedene, timecode-fähige Bandformate eingesetzt werden. Im Zweispur-Bereich waren dies zunächst die Pseudo-Video-Aufzeichnung auf U-matic und später DAT-Zuspeler und -Recorder, im Mehrspur-Bereich

DASH-Formate. Ende der 1980er Jahre kamen erste ausschließlich festplattenbasierte Schnittsysteme (Sonic Solutions, 1988) auf den Markt, seither gehören Schnitte und Überblendungen von beliebiger Dauer zu den Grundfunktionen digitaler Audioworkstations.

13.1.2 Praxis und Funktion

Beim analogen Schnitt auf Magnetband musste durch Hin- und Herbewegen des Bandes am Tonkopf (scrubbing) zunächst eine identifizierbare Stelle auf dem Band lokalisiert werden. Schnitte wurden daher überwiegend in oder kurz vor einem klar definierten Toneinsatz durchgeführt.

Digitale Schnittsysteme erlauben eine kontrolliertere Schnittbearbeitung durch die Möglichkeit, Dauer und Hüllkurve der Überblendung frei zu wählen und den Schnitt lediglich als Sprungbefehl beim Abspielen der Audiodaten zu definieren, ohne also das originale Audiomaterial zu verändern (non-destructive editing). Die Hüllkurve der Überblendung kann dabei so gewählt werden, dass der bei einem linearen crossfade (wie bei analogen Bandschnitten) unvermeidliche Pegel einbruch in der Mitte der Überblendung vermieden werden kann.

Bei einer linearen Ein- und Ausblendung der Amplitude fallen in der Mitte eines symmetrischen crossfades die -6 dB-Punkte der beiden Segmentenden zusammen

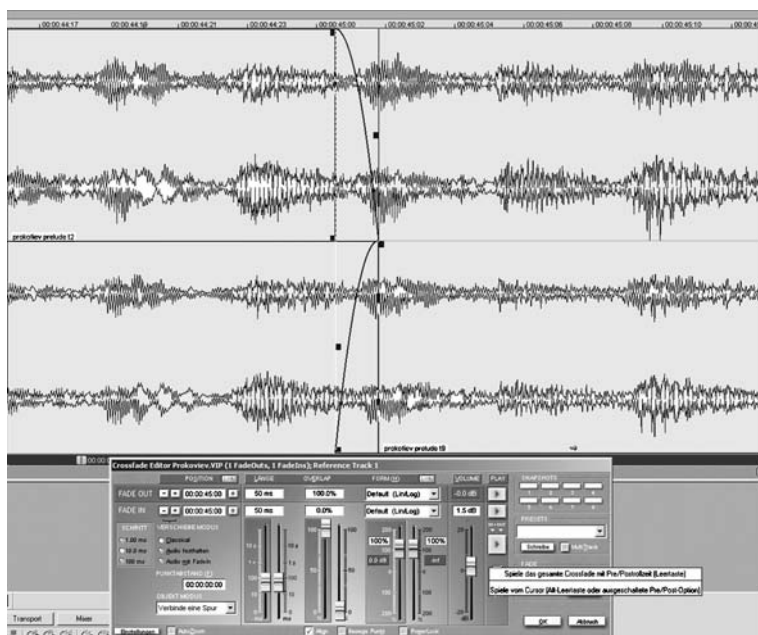


Abb. 13.1 Fenster zur Schnittbearbeitung mit einer Darstellung von Quell- und Zielspur (source/destination track) und einer Überblendung mit beliebiger Dauer und Hüllkurve

und addieren sich aufgrund der nicht-kohärenten Signalbeziehung zu einem Summenpegel von -3 dB. Digitale Audioworkstations bieten daher unter anderem eine Hüllkurve an, welche die Signalenergie während der Überblendung konstant hält. Seit der Einführung digitaler Audioworkstations kann auch die visuelle Darstellung der Wellenform als Orientierung für die Montage verwendet werden (Abb. 13.1). Als Konsequenz hat die Häufigkeit von Schnitten mit Einführung der Digitaltechnik im Mittel etwa um einen Faktor 2 bis 3 zugenommen (Weinzierl u. Franke 2002) und mittlere Segmentlängen von 10 bis 20 s zwischen zwei Schnitten dürften bei klassischen Musikproduktionen heute die Regel sein.

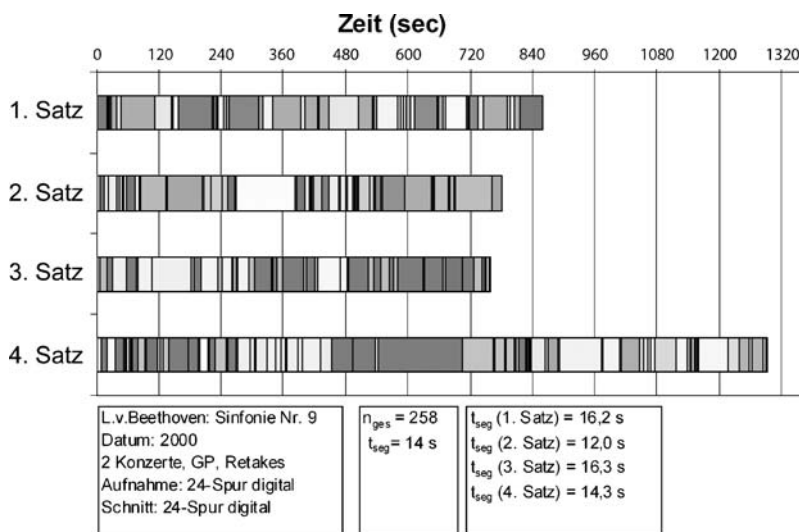


Abb. 13.2 Schnittsegmente und -zeitpunkte einer Produktion der Neunten Symphonie (Sätze 1–4) von L. v. Beethoven aus dem Jahr 2000, veröffentlicht bei einem Major-Label. Aufnahme und Schnitt erfolgten auf digitalem 24-Spur-Magnetband. Insgesamt wurden 267 Schnitte ausgeführt (Weinzierl u. Franke 2002).

Hinsichtlich der Funktion von Editierungsvorgängen kann unterschieden werden zwischen einem Insert-Schnitt, der als kurze Korrektur innerhalb eines längeren Takes das zeitliche Gefüge des korrigierten Abschnitts im Wesentlichen unangetastet lässt. Demgegenüber entspricht dem Assemble-Schnitt eine Überblendung von einer Aufführung des Werks in eine andere Aufführung, wobei zwangsläufig die ursprüngliche Zeitstruktur der Interpretation ersetzt wird durch die Illusion eines (im besten Fall) organischen, musikalischen Ablaufs. Die Begriffe stammen ursprünglich aus dem Umkopierschnitt, wo beim Insert-Schnitt das durch einen auf dem Band vorhandenen Timecode vorgegebene Zeitraster nicht verändert werden konnte, während beim Assemble-Schnitt auch der Timecode geschnitten wurde. Sie beschreiben im übertragenen Sinn auch die moderne Praxis, auch wenn bei festplattenbasierten, nicht-linearen Schnittsystemen keine Timecode- und Steuerspuren mehr geschnitten werden. Einen dritten Typus bilden zeitmanipulative Schnitte, bei denen – ohne in einen anderen Take zu wechseln – die Zeitstruktur der Aufnahme

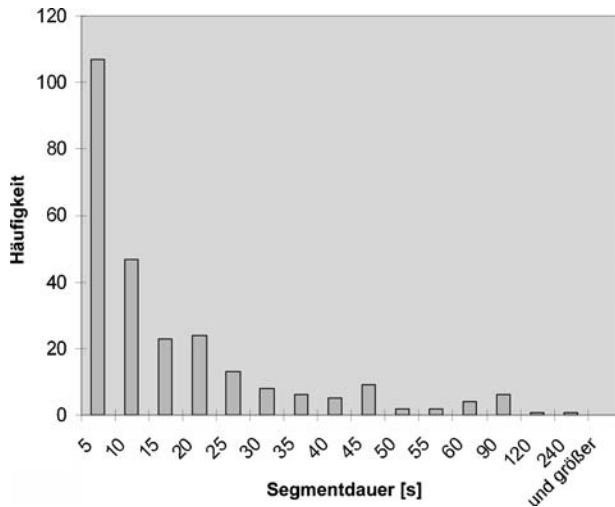


Abb. 13.3 Verteilung (Histogramm) der Segmentdauern aus Abb. 13.2.

verändert wird, etwa beim „Geradeschneiden“, wo der Einschwingvorgang bei mehrstimmigen Einsätzen soweit verkürzt wird, bis ursprünglich vorhandene Unsynchronitäten beim Akkordeinsatz nicht mehr hörbar sind. Zeitmanipulativ ist auch das Verkürzen oder Verlängern von Tonabständen und Pausen, wodurch die Zeitgestaltung der aufgezeichneten Aufführung direkt bearbeitet wird. Bei Algorithmen zur Time Compression oder zum Time Stretching geschieht dies auf Grundlage einer vorangehenden Signalanalyse automatisch (Abschn. 13.2.13 und Kap. 15.3.5).

13.2 Klanggestaltung

13.2.1 *Fader und Mute*

Mit dem Fader (Pegelsteller) wird die Amplitude des Audiosignals frequenzneutral verändert. Die Signalamplitude bestimmt in erster Linie die Lautstärke des Klangs, beeinflusst aber auch die wahrgenommene Entfernung und Größe der Schallquelle sowie – wegen der Kurven gleicher Lautstärke (Abb. 2.5) und aufgrund von Verdeckungseffekten bei der Mischung – ihre Klangfarbe. Der Fader wird sowohl im Hinblick auf eine korrekte Aussteuerung (Kap. 10.1.1) als auch mit dem Ziel einer künstlerisch-technischen Klangregie eingesetzt. Unter einer Vielzahl von technischen Ausführungen lassen sich folgende Realisierungsformen unterscheiden:

- Dreh- oder Schieberegler (sowie weitere, v. a. historische Bauformen),
- kontinuierlich oder stufenweise veränderbar,

- linear oder logarithmisch skaliert,
- Analoge Schaltungsdesigns T, H, L oder π (vgl. hierzu Webers 1999, 299–302),
- passiv (dämpfend) oder aktiv (nicht dämpfend),
- direkte oder spannungs- bzw. digital gesteuerte Signalverarbeitung,
- nur manuell steuerbar oder zusätzlich automatisierbar,
- nur manuell beweglich oder zusätzlich motorisiert,
- ohne oder mit Berührungssensor,
- physisch oder nur visuell durch Software repräsentiert.

In der professionellen Tonstudiotechnik ist heute der stufenlos veränderbare, automatisierbare, berührungsempfindliche und motorisierte Schieberegler Standard, der Steuerdaten für eine zumeist ausgelagerte digitale Signalverarbeitung liefert. Fragen der Skalierung oder analoger Schaltungsdesigns verlieren damit ihre Bedeutung. Nach wie vor wichtig bleiben jedoch ergonomische und funktionale Aspekte, denn bei dem Pegelsteller handelt es sich um das am häufigsten verwendete Audiobearbeitungsmittel.

Mutes (Stummschaltungen) stellen einen einfachen aber wirksamen Eingriff im Rahmen einer Audioproduktion dar: Signale werden unhörbar gemacht, indem ihr Pegel auf $L = -\infty$ reduziert wird. Zur Vermeidung von Knackgeräuschen geschieht dies entweder als Regelvorgang in einem kurzen Zeitintervall in der Größenordnung von 10 ms (meist bei zeitlinearen Produktionsweisen, etwa am Mischpult) oder als abrupter, jedoch auf den nächstliegenden Nulldurchgang des Audiosignals verschobener Schaltvorgang (meist bei nicht-zeitlinearen Produktionsweisen, z. B. in einem Harddisc-Recording-System). Je nach Möglichkeit und Zielsetzung wird das Mute per Hand zeitlich frei, durch Synchronisation zeitbezogen oder durch Marker präzise signalverlaufsbezogen ausgelöst. Das Mute kann nicht nur zur Entfernung von Störgeräuschen eingesetzt werden, es kann auch musikalisch-dramaturgischen Zwecken dienen: Durch das Stummschalten beliebig groß- oder kleinteiliger musikalischer Einheiten können Bestandteile eines Musikstücks entfernt werden, die das kompositorische Gefüge maßgeblich konstituieren, was im allgemeinen eine gravierende Änderung der musikalischen Wirkung zur Folge hat. Schließlich kann es auch zur Klanggestaltung eingesetzt werden: In einem Signalabschnitt, der als zusammenhängender Klang wahrgenommen wird, verursachen kurzzeitige Mutes als Hüllkurve der Signalamplitude einen ‚zerhackten‘ Klangverlauf, was z. B. bei der Produktion von Popularmusik als Effekt gewünscht sein kann.

13.2.2 *Panoramaregler*

Der Panoramaregler (auch Panoramapotiometer oder kurz Panpot) im Kanalzug von Mischpulten dient der Festlegung der Lokalisation eines Klangs in dem von den Lautsprechern aufgespannten Stereo-Panorama (Abb. 13.4). Er leitet das Signal mit einer durch seine Stellung definierten Pegeldifferenz auf die Ausgangswege des Mischpults, wodurch eine Phantomschallquelle zwischen den Wiedergabelautspre-

chern entsteht (Kap. 3.3.1). Der Zusammenhang zwischen (zweikanal-stereofoner) Pegeldifferenz und Hörereignisrichtung auf der Lautsprecherbasis wurde als Lokalisationskurve durch eine Reihe von Hörversuchen ermittelt (Abb. 10.15 oben).

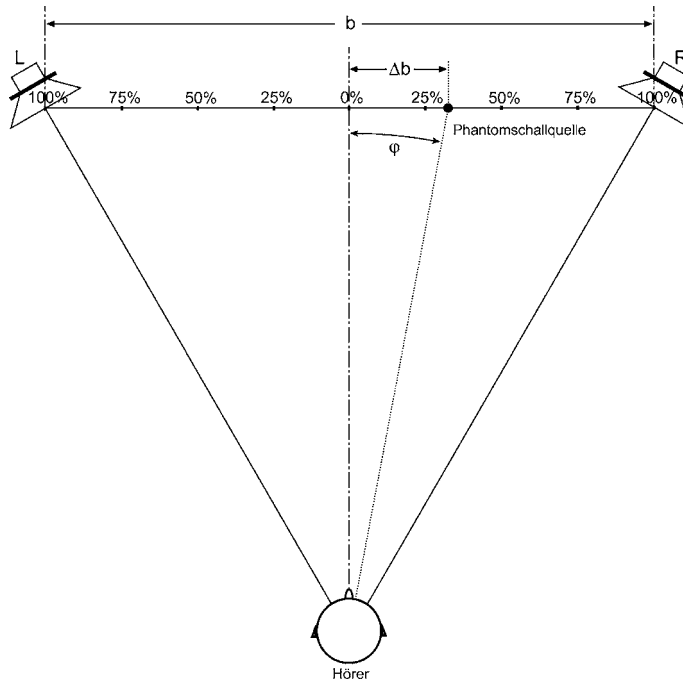


Abb. 13.4 Hörereignisrichtung in der Standard-Stereoanstellung (Kap. 11.1)

Da die stereofone Lokalisation nicht nur durch Pegel-, sondern u. a. auch durch Laufzeitdifferenzen beeinflusst werden kann, könnten beim Panoramaregler im Prinzip beide Gesetze genutzt werden. Der Einsatz von Laufzeitdifferenzen verringert jedoch die Monokompatibilität zweikanaliger Signale aufgrund von Kammfiltereffekten (Abschn. 13.2.10); daher ist die Erzeugung von Pegeldifferenzen das übliche Panorama-Konzept. Sie wird in analogen Systemen durch Aufmischung des Nutzsignals auf die Sammelschienen über gegensinnig variable Mischwiderstände realisiert oder digital durch eine Multiplikation der Amplitudenwerte des zweikanalig vervielfältigten Nutzsignals mit gegensinnig dimensionierten Faktoren.

Der Zusammenhang zwischen Reglerstellung und Kanaldämpfung variiert zwischen Mischpulten verschiedener Modelle und Hersteller (Abb. 13.5, Maempel 2001, S 173, 237). Unter der Voraussetzung, dass sich aufgrund der im weitgehend diffusen Schallfeld von Regieräumen unkorrelierten Ohrsignale beim Hörer die Leistungen und nicht die Schalldrücke der Lautsprecher addieren, müsste die Mittendämpfung des Panoramareglers 3 dB betragen. Dies ist bei den meisten Geräten der Fall, bestimmte Hersteller implementieren eine Mittendämpfung von 4,5 dB (Abb. 13.5).

Bei Audioproduktionen für Mehrkanal-Abhörformate (z. B. 5.x-Surround) sind mehrere Lautsprecherbasen vorhanden, allerdings ist die Lokalisation von Phan-

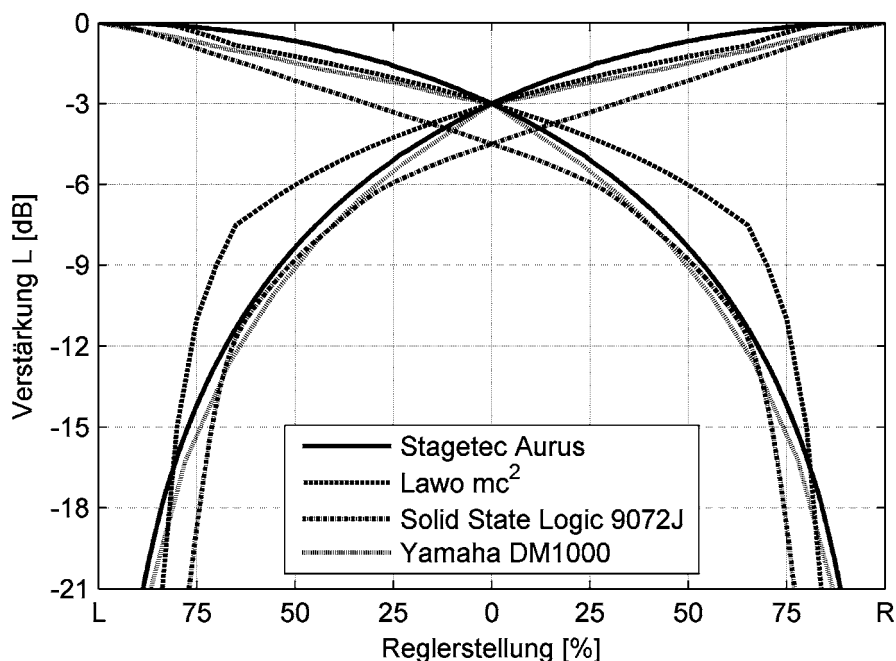


Abb. 13.5 Charakteristiken von Panoramareglern verschiedener Mischpulte (Stagetec, Lawo: Herstellerangaben; SSL, Yamaha: Messungen, z. T. geglättet)

tomschallquellen im hinteren und insbesondere im seitlichen Bereich wesentlich ungenauer und instabiler als im vorderen Bereich (Kap. 3.3.1).

Bei 5.x-Surround-Sound bestimmt der Panoramaregler für die Links-Rechts-Lokalisation im Vergleich zur Zweikanal-Stereofonie auch den Signalanteil für den Center-Kanal. Inwieweit mittig zu lokalisierende Signale tatsächlich durch den Mittenslautsprecher als Realschallquelle oder nach wie vor durch die Seitenlautsprecher als Phantomschallquelle dargestellt werden, kann mit dem zusätzlichen Parameter Divergenz geregelt werden. Ergänzend ist ggf. die grundsätzliche Mittendämpfung des Center-Kanals wählbar. Ferner ist ein weiterer Panoramaregler für die Lokalisationsrichtungen vorne/hinten vorhanden. Häufig besteht die Möglichkeit, die Pegelverhältnisse nicht nur über Drehregler, sondern auch über einen Joystick zu kontrollieren.

Neben Panoramakzepten, die auf der Erzeugung reiner Pegeldifferenzen beruhen, implementieren manche Hersteller weitere Verfahren zur Panorama-Einordnung aufnahmeseitig monophon kodierter Signale: Zeitdifferenzen, Pegel-Zeit-Differenz-Kombinationen, Spektraldifferenzen oder Crosstalk-Cancelling-Algorithmen. Ein als Virtual Surround Panning bezeichnetes Panoramakzept für Surround bezieht neben Laufzeitdifferenzen auch frühe Reflexionen ein, die je nach gewählter Lokalisation mittels Delays oder der Verrechnung mit Raumimpulsantworten generiert werden (Ledergerber 2002).

Speziell für Kopfhörerwiedergabe stehen Plug-ins zur Verfügung, die eine dreidimensionale Panoramaregelung weitgehend reflexionsfreier Audiosignale ermög-

lichen. Die binauralen Lokalisationscues werden dabei durch Verrechnung des Nutzsignals mit HRTFs (head-related transfer function) wählbarer Kunstköpfe erzeugt. Zusätzlich können frühe Reflexionen und Hall sowie ein Dopplereffekt bei Positionsänderungen generiert werden (Wave Arts 2005).

Surround-Panoramakonzepte werden häufig in Form einer punktförmig dargestellten Schallquelle auf der durch die Lautsprecher umstellten Fläche visualisiert. Dabei ist allerdings zu berücksichtigen, dass die Panning Laws (Gerzon 1992) 1. nur für den Sweet Spot (die mittige Abhörposition) gelten, 2. keine stabile seitliche Phantomschallquelle entstehen kann und 3. Lokalisationspositionen innerhalb der umstellten Fläche im Sinne einer Holophonie nicht realisierbar sind. Zur Weiterentwicklung von Panorama-Konzepten vgl. Neoran (2000) und Craven (2003).

Aufnahmeseitig bereits zweikanal-stereofon kodierte Audiosignale müssen bei einer Positionierung auch zweikanalig verarbeitet werden. Dies geschieht einerseits mit einem Balance-Regler, der die Pegel beider Kanäle gegensinnig variiert und dadurch die Lokalisation des Gesamtklangbilds verschiebt, andererseits mit einem Stereobreitenregler (width), der eine Einengung der Stereobreite durch die Zumischung der vertauschten Kanäle bewirkt oder eine Verbreiterung der Stereobreite durch die Zumischung der phaseninvertierten vertauschten Kanäle. In beiden Fällen wird eigentlich nur das Verhältnis von positiv bzw. negativ korrelierten und unkorrelierten Signalanteilen verändert, so dass es neben der gewünschten Wirkung zumeist auch zu Gewichtungverschiebungen im Gesamtklangbild kommt (s. Abschn. 13.2.3).

Grundsätzlich orientiert sich bei klassischer Musik und bei bildbezogenen Audioinhalten die Einordnung der Schallquellen weitgehend an den tatsächlichen bzw. gesehenen Schallquellenpositionen. Bei der Produktion von Popmusik hingegen werden Instrumente und Gesangsstimme häufig nach dem Kontrastprinzip gegenübergestellt: wichtige Monosignale mittig, sonstige Signale außen (vgl. Maempel 2001, 174–179) oder gelegentlich sogar hinten. Den größten Freiraum bei der Positionierung von Klangquellen bieten Hörspiel, Klangkunst und elektroakustische Musik. Grundsätzlich sollten jedoch ganz oder anteilig lauffzeitstereofon kodierte Signale außen positioniert werden, da eine Einengung mittels Panorama- oder Stereobreitenregler (Mischung) ungewünschte Kammfiltereffekte (vgl. 13.2.10) verursachen kann.

13.2.3 Stereo-Matrix

Die Stereo-Matrix (auch Richtungsmischer oder Summen-Differenzübertrager) dient der Umkodierung eines MS-Signals, das z. B. durch ein entsprechendes Mikrofonierungsverfahren erzeugt wird (Kap. 10.3.1.2), in ein zweikanal-stereofon kodierte Signal (XY). Dabei wird ein Ausgangssignal durch Summenbildung und ein Ausgangssignal durch Differenzbildung der Eingangssignale gewonnen, ggf. mit anschließender Dämpfung von 3 dB. Diese sog. Stereoumsetzung zwischen XY und MS ist nach dem gleichen Prinzip reversibel. Reine Stereo-Matrizen sind gemäß dieser Anwendung am häufigsten als Komponenten von Mischpulten anzutreffen. Einstellbare Parameter sind Abbildungsrichtung und Abbildungsbreite. Heute

weiter verbreitet ist ihr zweistufiger Einsatz als allgemeiner Stereobreiten- oder Width-Regler für zweikanal-stereofone Signale in Mischung und Pre-Mastering: Durch die MS-Kodierung eines solchen Signals, die Verschiebung des Amplitudenverhältnisses von M- und S-Signal und die anschließende Rückkodierung kann das Verhältnis von gleich- und gegenphasigen Anteilen in einem Stereosignal verändert werden. Diese Form der Bearbeitung wird vor allem zur Korrelationskorrektur von Signalen aus elektronischen Klangerzeugern, Effektgeräten oder fertigen Abmischungen eingesetzt. Mit der Korrelation hängt die Empfindung der Abbildungsbreite zusammen, auf die die Werte des Parameters Width Bezug nehmen (Abb. 13.4): Unter der Voraussetzung eines die Lautsprecherbasis ausfüllenden Klangbilds, das wenigstens partiell dekorreliert ist, führt danach ein Wert von 0 % (nur M-Signal) zu einem monofonen und mittigen Höreindruck, ein Wert von 100 % (gleiches Verhältnis von M- und S-Signal) zu einer unverändert vollständigen Klangbildausdehnung auf der Lautsprecherbasis und ein Wert von 200 % (nur S-Signal) zu einem Eindruck von Überbreite und fehlender Mitte, meist unter Verlust eindeutiger Lokalisation. Bei der Bearbeitung von Abmischungen geht mit der Korrelationsänderung eine Änderung des Mischungsverhältnisses zwischen mittig positionierten Monosignalen und gering korrelierenden außen positionierten Signalen einher. Da sich dadurch die Klangbilddance und der Raumeindruck verändern, werden Korrelationskorrekturen einer gesamten Abmischung, wenn überhaupt, nur geringfügig vorgenommen. Eine insbesondere im Pre-Mastering genutzte Variante ist die MS-Kodierung der Abmischung, die unterschiedliche Filterung oder Dynamikbearbeitung des M- und des S-Kanals und die anschließende Rückwandlung in ein XY-Signal, wodurch korrelationsbezogene Bearbeitungen möglich werden.

13.2.4 Kompressor

Allgemein bezeichnet man Verstärker, die in Abhängigkeit von der Spannung des Eingangssignals ihren Verstärkungsfaktor ändern, als Regelverstärker. Die Änderung kann gegensinnig erfolgen (Kompressor) oder gleichsinnig (Expander). Einstellbar sind neben der Ansprechschwelle und dem Verhältnis der Dynamikeinengung die Reaktionszeit und die Abklingzeit (sowie selten die Haltezeit). Bei der Wahl extremer Einstellungen werden der Kompressor als Limiter (Begrenzer) und der Expander als Noise Gate (Tor) bezeichnet. Kompressoren dienen der Verringerung, Expander der Vergrößerung der technischen Dynamik (vgl. terminologisch Jakubowski 1985), welche nicht mit der musikalischen Dynamik bzw. Spieldynamik zu verwechseln ist. Die Dynamikänderung wurde im analogen Zeitalter kombiniert in Kompander-Systemen mit dem Ziel der Rauschunterdrückung auf Übertragungstrecken oder Speichermedien eingesetzt (vgl. Webers 1989, S 323ff). Heute dient sie vor allem der Anpassung der Dynamik an die Erfordernisse von Übertragungswegen und der Klanggestaltung. Regelverstärker sind als Peripheriegeräte ausgelegt, als Bestandteil von Kanalzügen professioneller Mischpulte oder als Plug-ins für Audioworkstations.

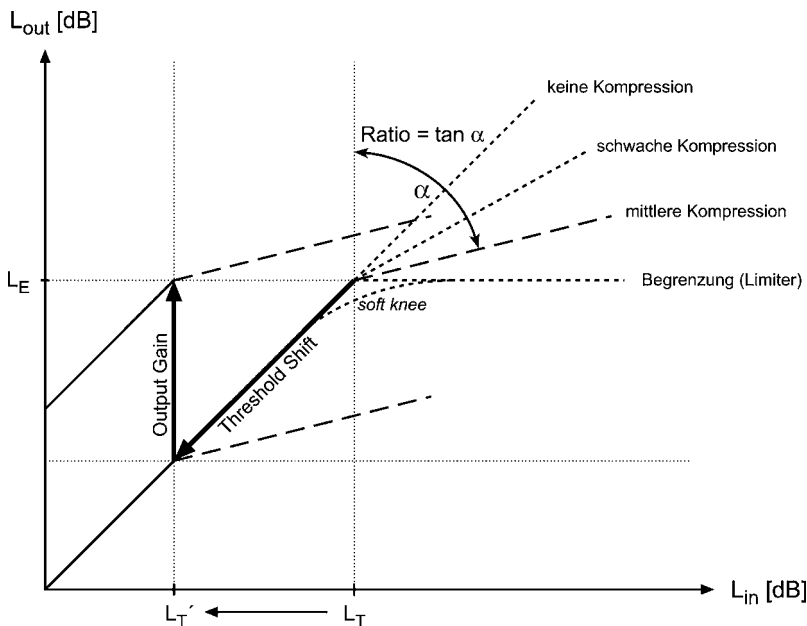


Abb. 13.6 Kennlinien des Kompressors

Ein Kompressor ist ein Verstärker, dessen Verstärkungsfaktor sich bei über einer definierbaren Schwelle (threshold) liegenden Eingangssignalen in einem vorbestimmt festen oder variablen Verhältnis (ratio) verringert. Dazu wird bei analoger Bauweise ein VCA (voltage controlled amplifier) von einem Regelkreis (side chain) gesteuert, der sich meist aus dem gleichgerichteten und integrierten Nutzsignal speist, wobei verschiedene Schaltungskonzepte möglich sind, z. B. die Verwendung des unregulierten Nutzsignals als Steuersignal (Vorwärtsregelung) oder des geregelten (Rückwärtsregelung). Digital wird die Kompression durch eine Multiplikation der Amplitudenwerte mit einem signalabhängigen Faktor realisiert. Mit der Option RMS/Peak kann bei vielen Geräten ausgewählt werden, ob der Kompressor auf den Effektivwert oder den Spitzenwert des Signals reagiert. Generell lässt sich zwischen dem statischen und dem dynamischen Verhalten unterscheiden.

13.2.4.1 Statisches Verhalten

Als statisches Verhalten bezeichnet man die Arbeitsweise des Kompressors in Abhängigkeit vom Pegel. In der Regel sind hierfür drei Parameter einstellbar, die aber in der Praxis nicht einheitlich benannt werden:

- Threshold = – Input Level = – Input Gain
- Ratio
- Output Gain = Output Level = Make Up Gain = Compression Gain = Hub

Die nichtlineare Kennlinie eines Kompressors setzt sich aus zwei Abschnitten zusammen, einem neutralen und einem flacheren, deren Übergang bei aktivierter Option Soft Knee geglättet ist (Abb. 13.6). Die Abflachung der in der Regel logarithmisch in dB dargestellten Kennlinie oberhalb der Ansprechschwelle hat zur Folge, dass eine Pegelerhöhung des Eingangssignals nicht mehr dieselbe, sondern nur eine geringere Pegelerhöhung des Ausgangssignals bewirkt. Sie wird durch den Parameter Ratio bestimmt, der als Verhältnis von Eingangspegeldifferenz ΔL_{in} und Ausgangspegeldifferenz ΔL_{out} oberhalb des Schwellwerts definiert ist:

$$R = \frac{\Delta L_{\text{in}}}{\Delta L_{\text{out}}} \quad (13.1)$$

R wird meist nicht als Zahlenwert, sondern als Verhältnis angegeben, z. B. $R = 6:1$, und kann mit

$$R = \tan \alpha \quad (13.2)$$

auch aus dem Abflachungswinkel α des oberen Kennlinienteils berechnet werden (Abb. 13.6).

Der Schwellwert L_T (also die Eingangsspannung, oberhalb derer die abgeflachte Kennlinie wirkt) wird mit dem Parameter Threshold bestimmt. Bei Annahme eines bestehenden Aussteuerungsziels am Kompressoreinsatzpunkt L_E gewinnt man durch das Absenken der Schwelle (Pfeil nach links unten und verschobene langgestrichelte Kennlinie) den Kompressionshub. Diesen sozusagen vergrößerten Headroom kann man nutzen, indem man das komprimierte Ausgangssignal mit dem Parameter Output Gain im Pegel anhebt (Pfeil nach oben und verschobene langgestrichelte Kennlinie). Dadurch wird bei gleichem Maximalpegel, also bei Einhaltung des Aussteuerungsziels, das Signal energiereicher, denn auch die unterhalb der Schwelle liegenden Signalabschnitte haben auf diese Weise einen höheren Ausgangspegel (verschobene durchgezogene Linie). Abb. 13.7 zeigt als Ergebnis einer solchen Audibearbeitung eine Zunahme des mittleren Pegels bzw. der akustischen Leistung und infolgedessen der Lautheit.

13.2.4.2 Dynamisches Verhalten

Als dynamisches Verhalten bezeichnet man die zeitabhängige Arbeitsweise des Kompressors. Sie ist bedingt durch das Verstreichen einer Reaktionszeit nach Über- und Unterschreitung des Schwellwerts. In diesem Zusammenhang sind folgende Parameter einstellbar, sie werden ebenfalls nicht immer einheitlich benannt:

- Attack
- Release = Recovery = Decay
- Hold (selten)

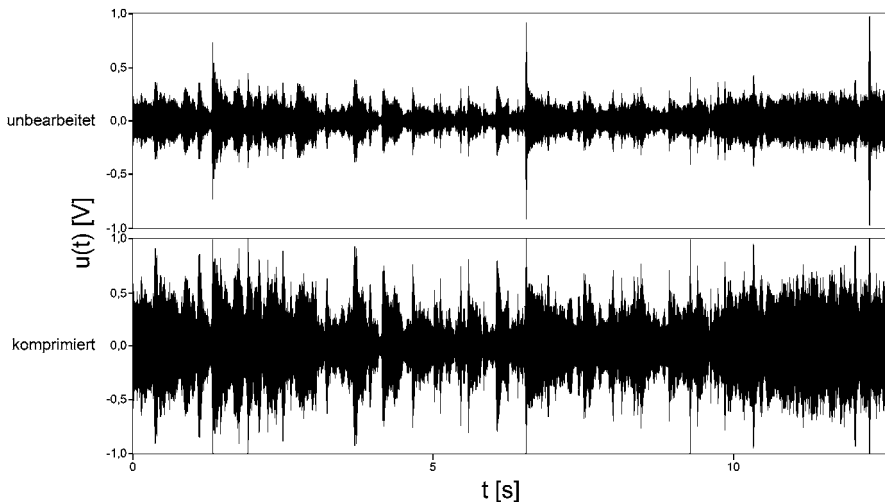


Abb. 13.7 Zwei Wellenformdarstellungen desselben Musikausschnitts. Das komprimierte Signal (unten) besitzt bei gleichem Maximalpegel eine höhere akustische Leistung und klingt lauter.

Die Zeitkonstante Attack ist ein Maß für die Zeitspanne, die der Kompressor für die Reduktion des Verstärkungsfaktors benötigt, nachdem das Eingangssignal den Schwellwert überschritten hat (Ansprechzeit). Release ist ein Maß für die Zeitspanne, die der Kompressor für die Rückführung der Verstärkung auf den Faktor 1 nach dem Sinken des Eingangssignals unter den Schwellwert benötigt (Rückstellzeit). Als Attack-Zeit Δt_{att} ist diejenige Zeitspanne definiert, die nach plötzlicher Überschreitung des Schwellwerts um 10 dB vergeht, bis 63 % ($= 1 - 1/e$) der Differenz zwischen der neuen Eingangsspannung L_{in} und der neuen ratioabhängigen Ausgangsspannung $L_{\text{out}}(t_2)$ ausgeregelt wurden (Abb. 13.8). Die Release-Zeit Δt_{rel} ist demgegenüber die Zeit, die nach einem plötzlichen Eingangssignalabfall von 10 dB oberhalb der Threshold-Spannung auf die Threshold-Spannung L_T bis zu dem Zeitpunkt vergeht, an dem etwa 63 % der Differenz zwischen der noch ratioabhängigen Ausgangsspannung $L_{\text{out}}(t_1)$ und der neuen Ausgangsspannung $L_{\text{out}}(t_2)$ ausgeregelt wurden (Abb. 13.9).

Es ist leicht ersichtlich, dass ein Kompressor aufgrund des exponentiellen Verlaufs des Regelvorgangs in jedem Falle länger für das tatsächliche Ausregeln benötigt, als mit den Zeitkonstanten Attack und Release eingestellt wurde. Attack-Zeiten bewegen sich meistens im Bereich zwischen 50 μs und 50 ms, Release-Zeiten zwischen 10 ms und 3 s.

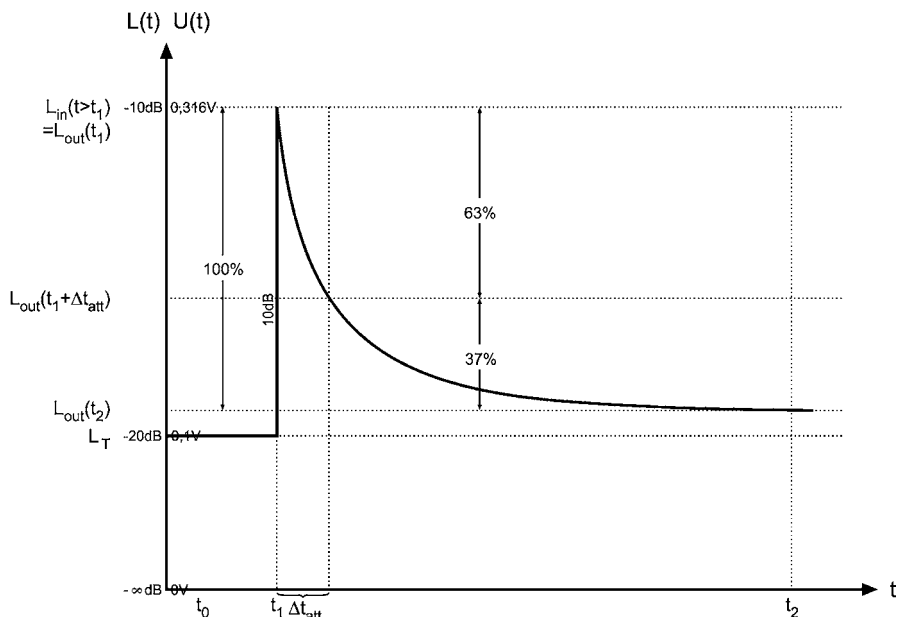


Abb. 13.8 Dynamisches Verhalten des Kompressors: Attack-Vorgang

13.2.4.3 Physikalische Signalveränderung

Die Auswirkungen verschiedener Ansprechzeiten auf hoch- und tieffrequente Signale zeigt Abb. 13.10. Dargestellt sind jeweils ein sinusförmiges Signal, dessen zunächst geringe Amplitude plötzlich über den Schwellwert (gepunktete Linie) steigt und durch den so verursachten Verstärkungsrückgang (gestrichelte Hüllkurve) komprimiert wird, sowie zum Vergleich das unkomprimierte Signal (Strichpunkt). Demnach gilt für das Ansprechen eines Kompressors: Ist ein langes Attack gewählt, bleiben Transienten erhalten (oben links), und es sind kurzzeitige Übersteuerungen der nachfolgenden Übertragungskette möglich, sofern kein Headroom mehr vorhanden ist. Ist ein kurzes Attack gewählt, werden Transienten wirksamer abgefangen (oben rechts), es erfolgt aber eine starke Deformation des Nutzsignals im tieffrequenten Bereich (unten rechts). Durch diese nichtlinearen Verzerrungen erhöht sich der Klirrfaktor. Beide Einstellungen haben also bestimmte Nachteile. Für den Rücklauf des Kompressors gilt: Ist ein langes Release gewählt, regeln kurze Spitzen für längere Zeit das nachfolgende Signal und bestimmen damit stark den mittleren Pegel. Die Wahl eines kurzen Release führt hingegen zu häufigen Regelvorgängen (Tabelle 13.1).

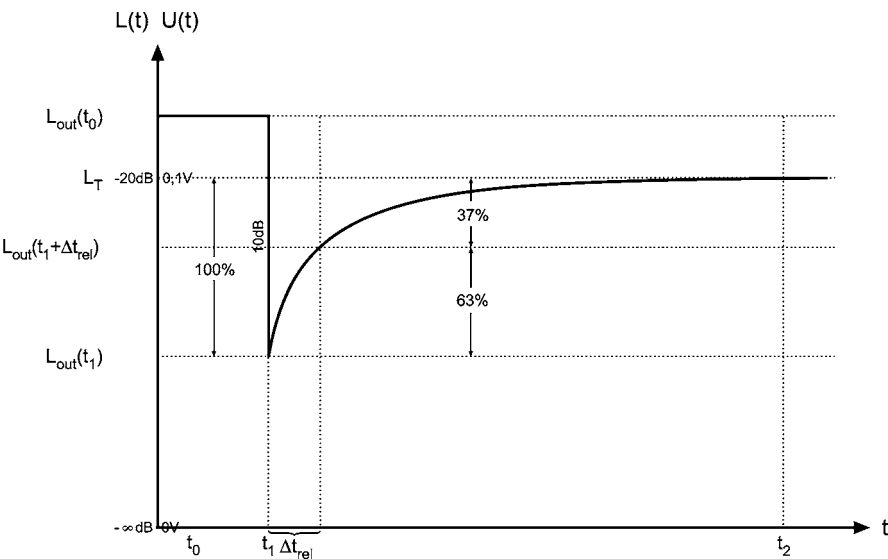


Abb. 13.9 Dynamisches Verhalten des Kompressors: Release-Vorgang

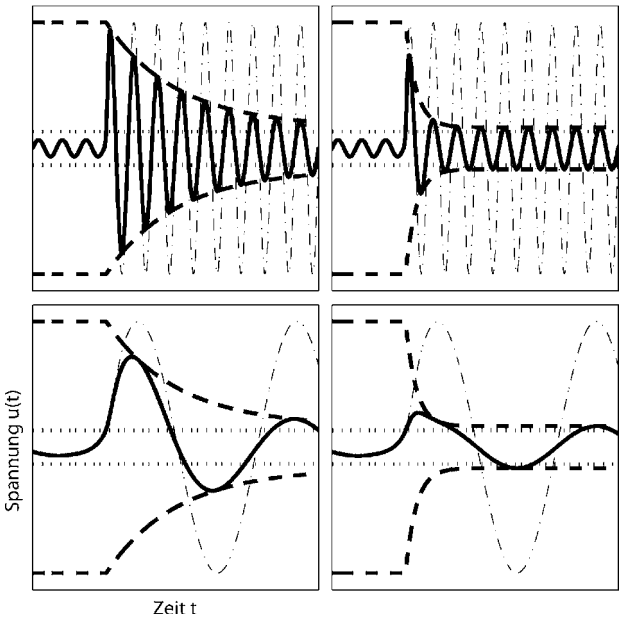


Abb. 13.10 Dynamisches Verhalten des Kompressors: Signaldeformation durch Attack-Vorgang

13.2.4.4 Perzeptive Wirkung

Im Hinblick auf die Wahrnehmung der durch den Kompressor bewirkten Amplitudenänderungen und nichtlinearen Verzerrungen ist das statische Verhalten für die Nutz- und das dynamische Verhalten überwiegend für die Störeffekte verantwortlich (Tabelle 13.1). Generell führen kleine Zeitkonstanten zu einer wahrnehmbaren Klangverdichtung, die gestalterisch beabsichtigt sein kann (z. B. in der Popmusikproduktion), jedoch auch die Lästigkeit des Gehörten erhöhen kann (vgl. Wagner 1997). Sofern nicht maximale Lautheit gefordert ist, lässt sich in der Praxis das Dilemma der Stör- und Nutzeffekte zufriedenstellend verringern, indem man zwei Kompressoren hintereinander einsetzt: Für jedes zu bearbeitende Einzelsignal einen Kompressor mit geringer Schwelle, kleiner Ratio und mittleren oder großen Zeitkonstanten für eine unauffällige aber effektive Dynamikreduzierung; und einen nur selten ansprechenden Kompressor mit hoher Schwelle, großer Ratio und geringen Zeitkonstanten bei gemischten Signalen für den rein technischen Übersteuerungsschutz, etwa an Summenausgängen. Eine Perfektionierung dieses Ansatzes bietet die signalabhängige Dynamisierung von Parametern (Abschn. 13.2.4.5). Zur klanglich optimalen Wahl von Kompressionsparametern bei verschiedenen Abhörbedingungen vgl. Wagenaars, Houtsma & van Lieshout (1986), zum Einsatz in der Hörgeräteakustik vgl. Neumann (1998).

13.2.4.5 Varianten

Verschiedene schaltungstechnische Abwandlungen des reinen Kompressors können effektivere, artefaktreichere, artefaktärmere, kreativere oder zuverlässigere Dynamikbearbeitungen ermöglichen.

Stereo- und Mehrkanalverkopplung: Die Side Chains zweier Kompressoren können elektrisch verbunden werden (stereo link). Dies führt dazu, dass Regelvorgänge, gleich durch welches Nutzsignal ausgelöst, von beiden Kompressoren ausgeführt werden, die in der Regel identisch eingestellt sind. Eine Verkopplung ist für die Kompression von aufnahmeseitig stereofon bzw. mehrkanalig kodierten Signalen erforderlich, wenn sich die Lautstärkebalance und Lokalisation im Klangbild nicht durch einkanalige Regelvorgänge verschieben soll.

Kompression verzögerter Signale: Das oben beschriebene Dilemma bei der Wahl geeigneter Zeitkonstanten, das sich entweder in hörbaren Artefakten oder dem nur unzuverlässigen Abfangen von Signalspitzen äußert, kann umgangen werden, wenn man das Nutzsignal gegenüber dem Steuersignal geringfügig verzögert. Ist die Verzögerung so groß wie die Attack-Zeit des Kompressors, weist das komprimierte Signal keine Transienten und kaum reaktionszeitbedingte Artefakte mehr auf. Dieses Prinzip findet sich bei sog. Transienten-Limitern sowie optional bei Mastering-Prozessoren. Laufzeitglieder sind digital leicht zu realisieren, und in digitalen Kompressoren ist die Funktion (predict, look ahead) häufig vorgesehen.

Tabelle 13.1 Physikalische und perzeptive Wirkungen des Dynamikkompessors

	Physikalische Wirkungen*	Perzeptive Wirkungen*	Nutzeffekt (+) Störeffekt (-)
Attack lang ($\Delta t_{\text{att}} \geq 3 \text{ ms}$)	Langsames Ausregeln. Kurze Spitzen werden nicht abgeregelt.	Unauffällige Dynamikverminderung; Lautheitserhöhung bei Nutzung des Kompressionshubs nicht maximal möglich. Verzerrungen durch Übersteuerung der nachfolgenden Übertragungskette möglich.	+ –
Attack kurz ($\Delta t_{\text{att}} < 3 \text{ ms}$)	Schnelles Ausregeln. Kurze Spitzen werden abgeregelt; geringere Impulshaftigkeit. Signaldeformation im tieffrequenten Bereich; nichtlineare Verzerrung; erhöhter Klirrfaktor.	Auffällige Dynamikverminderung; Signalspitzen bestimmen die Lautstärke; Klangverdichtung; Lautheitserhöhung bei Nutzung des Kompressionshubs maximal möglich. Unter Umständen größeres Empfinden von Weichheit (z. B. bei perkussiven Klängen). Verzerrungen bis hin zu knackähnlichen Störungen.	– (Gestaltungsmittel) + –
Release lang ($\Delta t_{\text{rel}} \geq 1 \text{ s}$)	Seltenes Regeln. Geringer mittlerer Pegel; bei kurzem Attack regeln kurze Signalspitzen nachhaltig den Pegel.	Unauffällige Dynamikverminderung; Lautheitserhöhung bei Nutzung des Kompressionshubs nicht maximal möglich. Geringe mittlere Lautstärke; bei kurzem Attack sind plötzliche nachhaltige Lautstärkeeinbrüche möglich.	+ –
Release kurz ($\Delta t_{\text{rel}} < 1 \text{ s}$)	Durch häufiges Regeln Modulation von langsam veränderlichen Signalanteilen.	Erhöhung von Rauigkeit oder Pumpen; Klangverdichtung; Lautheitserhöhung bei Nutzung des Kompressionshubs maximal möglich.	– (Gestaltungsmittel)

* Stärke und Häufigkeit der Wirkungen hängen außerdem von den Eigenschaften des Nutzsignals und dem gewählten Threshold-Wert ab.

Frequenzabhängige Kompression: Das Steuersignal kann gefiltert werden, entweder durch ein im Kompressor vorhandenes Filter oder das Einschleifen eines externen Filters. Am häufigsten wird diese Möglichkeit angewendet, indem mittels Equalizer, Hochpassfilter oder Bandpass hochfrequente Signalanteile relativ zu den restlichen verstärkt werden. Diese Signalanteile, die z. B. bei Sprache Artikulations- und Zischlaute sind, bestimmen so hauptsächlich den Regelvorgang des Kompressors, der sich auf das gesamte Nutzsignal auswirkt. Da bei Sprache diese Laute kaum gleichzeitig mit anderen auftreten, werden vor allem erstere vermindert, ohne dass das vorangehende oder nachfolgende Signal komprimiert würde. Auf diese Funktion spezialisierte Geräte, die z. B. zur Unterdrückung von S-Lauten eingesetzt werden, heißen De-Esser. Sie verfügen meist über einen Bandpass mit einer varia-

blen Eckfrequenz von 0,8 bis 8 kHz. Man kann einen solchen Vorgang so stark einstellen, dass die bearbeitete Stimme regelrecht lispelt.

Fremdansteuerung: Durch Speisung des Side Chains mit einem völlig anderen (meist kürzeren, repetitiven) Signal kann der im Pop- und Dance-Musik-Bereich beliebte Ducking-Effekt erzeugt werden: Z. B. wird ein stehender Klang im Rhythmus des Schlagzeugs im Pegel vermindert, so dass ein gezieltes Pumpen entsteht.

Multibandkompression: Das Nutzsignal wird in (meist drei bis fünf) Bänder zerlegt. Die Bänder werden zeitgleich in verschiedenen Kompressoreinheiten komprimiert und schließlich wieder zusammengemischt. Für verschiedene Frequenzbereiche lassen sich so unterschiedliche Kompressionsparameter einstellen, was in der Praxis meist mit dem Ziel der Leistungs- und damit auch Lautheitsmaximierung eingesetzt wird. Dies spielt insbesondere im Bereich des Rundfunks eine große Rolle, wo in der Regel von einem Zusammenhang zwischen Lautheit und Senderpräferenz ausgegangen wird, auch wenn es hierfür bislang keine empirischen Belege gibt. Man kann jedoch vermuten, dass diese besonders effektive Art der Klangverdichtung zumindest mittelfristig zu einer höheren Lästigkeit des Gehörten führt (Wagner 1997). Zurückhaltend eingesetzt kann ein Multibandkompressor hingegen einer Produktion den letzten Schliff verleihen.

Programmabhängige Parametersteuerung: Die Kompressionsparameter können unter Berücksichtigung psychoakustischer Zusammenhänge dynamisch an die Signaleigenschaften angepasst werden (Hartwich et al. 1995). Speziell für Payout-Wege stehen hochwertige Dynamikprozessoren zur Verfügung, die Spitzenwert- und Effektivwertmessung, eine Analyse des Signalverlaufs mit dynamischer Anpassung der statischen und dynamischen Parameter, eine Verzögerung des Nutzsignals und seriell geschaltete Regelstufen (Multi-Loop-Architektur) funktionell verbinden. Dadurch wird eine korrekte und zuverlässige Aussteuerung bei gleichzeitiger perzeptiver Unauffälligkeit der Regelvorgänge und hoher klangfarblicher Treue möglich. Diese Methode ist weniger für die Lautheitsmaximierung als vielmehr für eine normengerechte und klangqualitätserhaltende Modulationsaufbereitung von Sendesummensignalen geeignet.

13.2.4.6 Einsatzgebiete

Dynamikkompressoren können in allen Produktionsschritten eingesetzt werden:

- Bei der Aufnahme soll ein Kompressor die Dynamik hochdynamischer Signale (z. B. Pop- und Rockgesang oder Big-Band-Blechbläser) reduzieren, um sie zuverlässiger und höher aussteuerbar zu machen. De-Esser sollen hochfrequente Signalanteile gezielt reduzieren, um eine hohe Aussteuerbarkeit des aufzunehmenden Signals auf analogen Magnetbändern bei geringen Artefakten zu ermöglichen.
- In der Abmischung werden einzelne Spuren mit dem Ziel der Dynamikanpassung, der Klangverdichtung und dem „Hüllkurvendesign“ (z. B. bei perkussiven Signalen wie Trommeln), also aus gestalterisch-ästhetischen Gründen, kompri-

miert. Auch lässt sich mit dem Mittel der Fremdansteuerung kreativ oder experimentell arbeiten.

- Der Gesamtmix (das Summensignal) wird teilweise schon während der Abmischung komprimiert, mit dem Ziel der Klangverdichtung, eines ducking-ähnlichen Effekts oder der Lautheitserhöhung.
- Beim Pre-Mastering kommt v.a. Multibandkompression zum Einsatz, hauptsächlich mit dem Ziel der Lautheitserhöhung.
- Im Rundfunkstudio ist die Multibandkompression fester Bestandteil des wellentypischen Soundprocessings, das ein konstantes, wiedererkennbares Klangprofil erzeugen soll.
- Ausspiel- bzw. Sendewege werden durch Kompressoren vor technischen Übersteuerungen geschützt, sei es durch einen Begrenzer vor dem Plattenschneidstichel, durch einen Transientenlimiter vor Sende-Leistungsverstärkern (z. B. Senderöhren) oder Uplink-Satelliten.
- Selbst in Consumer-Geräten können Kompressoren enthalten sein, die optional für eine situationsgerechte geringe Programmdynamik sorgen sollen, z. B. beim Autofahren.

Die Auflistung zeigt, dass elektronische Medien heute fast ausschließlich komprimierte Tonsignale übertragen. Lediglich digitale Tonträger mit klassischer Musik werden nicht oder – bei Programminhalten mit sehr hoher Dynamik (spätromantische Orchestermusik, Orgelmusik) – nur geringfügig komprimiert.

13.2.5 Expander

Ein Expander ist ein Verstärker, dessen Verstärkungsfaktor sich bei einem unter einer definierbaren Schwelle (Threshold) liegenden Bereich von Eingangssignalen in einem vorbestimmt festen oder variablen Verhältnis (Ratio) vergrößert. Das Prinzip der technischen Realisation entspricht dem eines umgekehrten Dynamikkompessors. Demnach wird zwischen zu bearbeitendem Signal und Steuersignal unterschieden. Der Steuereingang wird hier auch Key Input genannt. Auch beim Expander kann zwischen statischem und dynamischem Verhalten unterschieden werden.

13.2.5.1 Statisches Verhalten

Als statisches Verhalten bezeichnet man die Arbeitsweise des Expanders in Abhängigkeit vom Pegel. Meistens sind hierfür drei Parameter einstellbar, die in der Praxis nicht einheitlich benannt werden:

- Threshold = – Input Level = – Input Gain
- Ratio (entfällt bei einigen Modellen)
- Range

Die Kennlinie des Expanders setzt sich aus drei Abschnitten zusammen, zwei neutralen und einem dazwischenliegenden steileren (Abb. 13.11). Die höhere Steigung der Kennlinie unterhalb der Ansprechschwelle L_T hat zur Folge, dass eine Pegelverminderung des Eingangssignals ΔL_{in} nicht mehr dieselbe, sondern eine stärkere Pegelverminderung des Ausgangssignals ΔL_{out} bewirkt. Die Steigung dieses Kennlinienabschnitts wird durch den Parameter Ratio bestimmt. Analog zum Kompressor ist die Ratio R definiert als das Verhältnis von Eingangs- und Ausgangspegeldifferenz (Formel 13.1) und kann über den Abflachungswinkel α berechnet werden (Formel 13.2, Abb. 13.11).

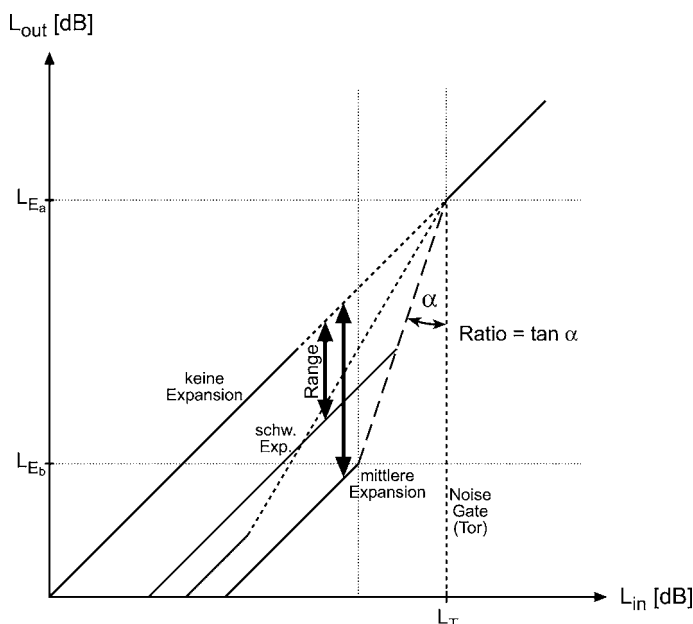


Abb. 13.11 Kennlinien des Expanders

Der Schwellwert oder Expandereinsatzpunkt (also die Eingangsspannung, unterhalb derer der steile Kennlinienabschnitt wirkt) wird mit dem Parameter Threshold bestimmt. Die Dämpfung von Signalen, deren Pegel in den unteren Kennlinienbereich fallen, wird mit dem Parameter Range in dB angegeben. Geht man davon aus, dass leise Signale Störsignale sind (z. B. Übersprechen bei Mikrofonaufnahmen, Hintergrundgeräusche), so kann mit ihm der Abstand des Nutzsignals zu diesem Störsignal dosiert vergrößert werden – eine Nichtüberschneidung beider Dynamikbereiche vorausgesetzt. Bei der Einstellung des Expanders als Noise Gate besteht die Kennlinie nur aus zwei Abschnitten, dem oberen neutralen und dem senkrecht verlaufenden mit der Ratio 0. In diesem Falle werden keine Signale unterhalb des Schwellwerts übertragen.

31.2.5.2 Dynamisches Verhalten

Wie beim Kompressor bezeichnet man auch beim Expander die zeitabhängige Arbeitsweise als dynamisches Verhalten. Sie ist bedingt durch das Verstreichen einer Reaktionszeit nach Über- und Unterschreitung des Schwellwerts. In diesem Zusammenhang sind folgende, nicht immer einheitlich benannte Parameter einstellbar:

- Attack
- Release = Recovery = Decay
- Hold

Die Zeitkonstante Attack ist ein Maß für die Zeitspanne, die der Expander für das Erreichen der neutralen Verstärkung benötigt, nachdem das Eingangssignal den Schwellwert überschritten hat (Ansprechzeit). Release ist ein Maß für die Zeitspanne, die der Expander für die Dämpfung des Signals benötigt, nachdem es unter den Schwellwert gesunken ist (Rückstellzeit). Häufig kann an Expandern der Release-Vorgang verzögert werden, um ein ungedämpftes Ausklingen des Nutzsignals zu gewährleisten. Diese Verzögerung oder Mindesthaltezeit wird mit dem Parameter Hold bestimmt. Das Zeitverhalten eines Expanders weist wie das eines Kompressors einen exponentiellen Verlauf auf, so dass die eigentlichen Regelvorgänge länger dauern, als durch die Maße Attack und Release angegeben. Attack-Zeiten bewegen sich meistens im Bereich zwischen 10 μ s und 20 ms, Hold-Werte zwischen 0 und 10 s, Release-Zeiten zwischen 50 ms und 10 s.

13.2.5.3 Physikalische Signalveränderung

Die Regelvorgänge des Expanders erstrecken sich über einen weiten Dynamikbereich. Insoweit kann es bei kurzen Attack-Zeiten zu starken Signaldeformationen kommen.

13.2.5.4 Perzeptive Wirkung

Im Gegensatz zum Kompressor ist die Zielsetzung beim Einsatz eines Expanders meistens, das Signal einer akustischen Quelle von Signalen anderer akustischen Quellen zu trennen, so dass nicht nur die Dynamik und Lautstärke des Klangs beeinflusst werden, sondern auch der Ein- und Ausklingvorgang. Neben dem zumeist großen Regelbereich macht diese Nähe der Regelvorgänge zu den wahrnehmungspsychologisch wichtigen Onset- und Offset-Cues eine perzeptiv unauffällige Einstellung der Expander-Parameter schwierig. Typische Artefakte sind Knackgeräusche oder An- und Abschnitte des Nutzsignals bzw. des Störsignals. Angemessen justiert kann ein Expander den empfundenen Störgeräuschabstand vergrößern. In den Übertragungswegen verschiedener Schallquellen eingesetzt, kann so die Durchhörbarkeit eines Klangbilds und die Prägnanz der einzelnen Schallquellen vergrößert werden. Typische Beispiele sind Diskussionsrunden und Schlagzeugaufnahmen.

13.2.5.5 Varianten

Die möglichen oder gebräuchlichen Abwandlungen des Einsatzes von Expandern sind im Grunde dieselben wie bei Kompressoren. Zur Bearbeitung von aufnahmeseitig zweikanal-stereofon oder mehrkanalig kodierten Signalen muss der Expander verkoppelt betrieben werden, indem die Steuereingänge zusammengeschaltet werden. Zur Vermeidung von Anschnitten des Nutzsignals durch die Attackzeit ist dessen Verzögerung gegenüber dem Steuersignal sinnvoll. In Live-Situationen ist dies nicht praktikabel, so dass nach Möglichkeit ein voreilendes Signal als Steuersignal verwendet wird, z. B. von Kontaktmikrofonen bei der Übertragung von Trommeln. Prinzipiell ist natürlich auch eine frequenzabhängige Ansteuerung eines Expanders möglich, wenngleich selten sinnvoll. Häufig angewandt wird eine Fremdansteuerung durch andere, insbesondere impulshafte Signale. So lässt sich etwa ein durchgängiger Klang durch den Einsatz eines fremdgesteuerten Noise Gates im Rhythmus eines perkussiven Sounds zerhacken (Gater-Effekt).

13.2.5.6 Einsatzgebiete

Expander werden in den Produktionsschritten Aufnahme, Abmischung, Pre-Mastering und Programmausspielung eingesetzt und dienen je nachdem der Störgeräuschverminderung oder der Klanggestaltung. In Dynamikprozessoren für Sendewege oder Masteringsysteme können die Parameter von Expandern auch programmabhängig steuerbar sein mit dem Ziel einer möglichst effektiven, artefaktfreien und zuverlässigen Störgeräuschverminderung.

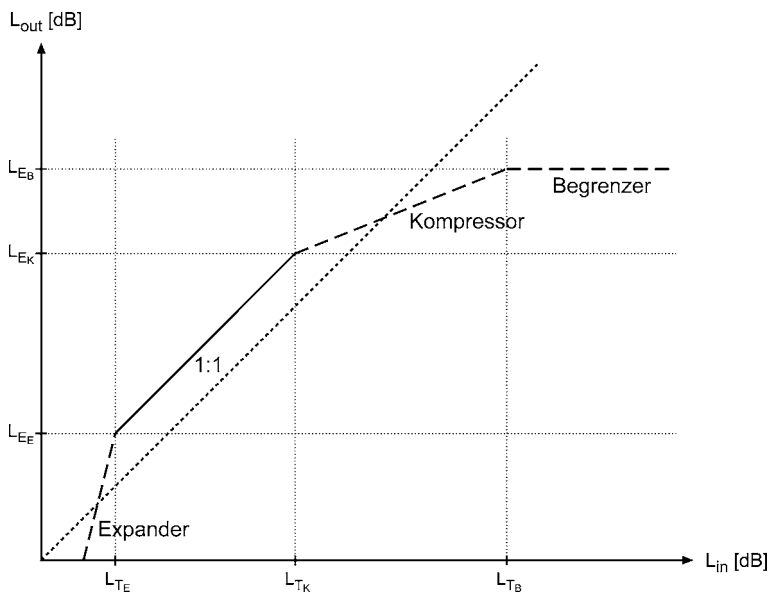


Abb. 13.12 Kennlinie eines kombinierten Regelverstärkers

13.2.6 Kombinierte Regelverstärker

Kompressoren und Expander können auch in Serie zu einem kombinierten Regelverstärker verschaltet werden, um die Handhabung der Dynamik zu vereinfachen. Dem entsprechend werden kombinierte Regelverstärker z. B. in weitgehend automatischen Übertragungs- oder Beschallungssystemen eingesetzt. In der gängigsten Kombination ergeben sich vier Kennlinienabschnitte und drei Schwellwerte für den Einsatz von Expander, Kompressor und Begrenzer (Abb. 13.12).

Kombinierte Regelverstärker sind Bestandteil moderner Dynamikprozessoren. Sie werden eingesetzt, um die Dynamik von Audioprogrammen an menschliche Dynamiktoleranzen anzupassen. Nach Lund (2006) lässt sich der tolerierbare Dynamikbereich durch drei Werte beschreiben. Einen Schwerpunkt: den der bevorzugten oder definierten Programmlautheit äquivalenten mittleren Abhörschalldruckpegel; eine Obergrenze: den präferierten oder systemtechnisch gegebenen maximalen Schalldruckpegel (Spitzenwert); und eine Untergrenze: den der Lautheit der Störgeräusche äquivalenten mittleren Schalldruckpegel. Die so für verschiedene Wiedergabesysteme bzw. -situationen ermittelten Ober- und Untergrenzen und die sich daraus ergebenden tolerierbaren Dynamikbereiche zeigt Abb. 13.13.

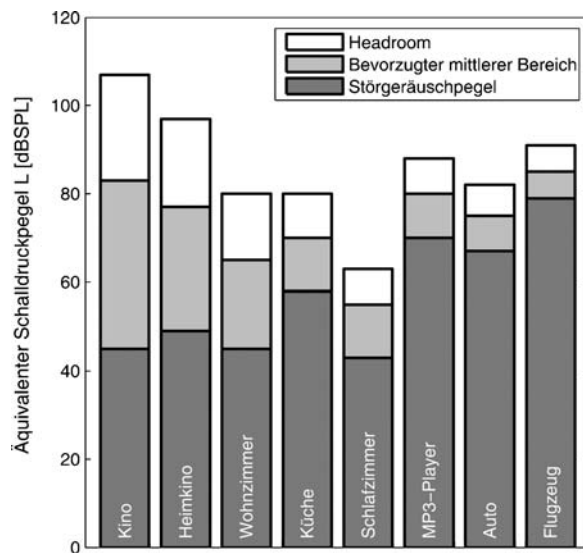


Abb. 13.13 Ober- und Untergrenzen tolerierbarer Dynamikbereiche von Audio-Konsumenten in verschiedenen Hörsituationen, dargestellt als Differenz zu den jeweils definierten oder empirisch präferierten lautheitsäquivalenten Wiedergabeschalldruckpegeln (z. B. Kino: 83 dB SPL), nach (Lund 2006)

13.2.7 Filter

Nach etablierter Audio-Terminologie dienen Filter der klangfarblichen Kontrolle von Audiosignalen durch Veränderung des Frequenzspektrums von Signalen. Für die klangfarbliche Wirkung ist dabei v.a. der Amplitudengang von Bedeutung, der Phasengang hingegen ist in bestimmten Grenzen vernachlässigbar. Technisch sind Filter analog als passive oder aktive RC-Schaltungen konzipiert, in letzterem Falle z. B. als Operationsverstärker mit frequenzabhängiger Gegenkopplung. In der digitalen Signalverarbeitung werden generell Algorithmen zur Addition und Multiplikation von Abtastwerten als Filter bezeichnet, selbst wenn der Amplitudengang dabei nicht verändert wird (Kap. 15.3.1). Filter sind als Mischpult-Komponenten, Stand-alone-Hardware oder Plug-ins verfügbar.

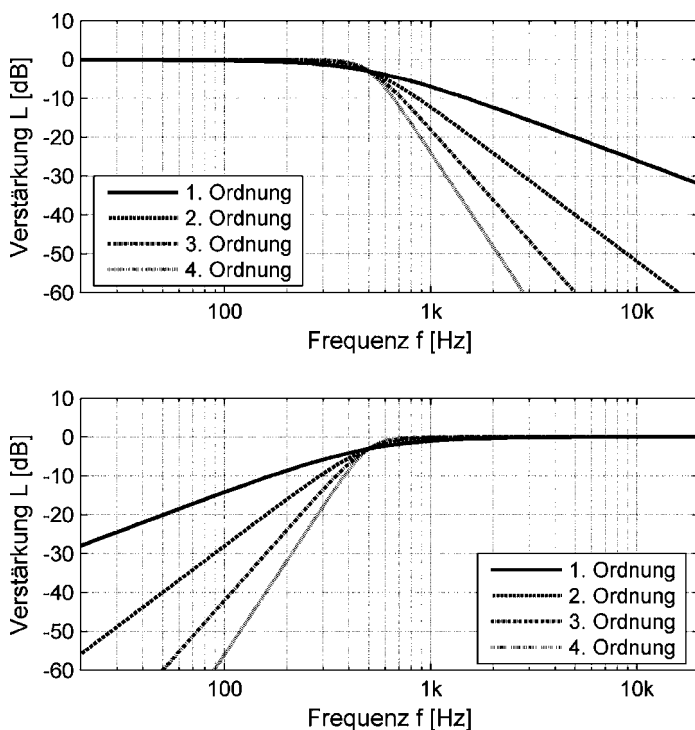


Abb. 13.14 Tiefpassfilter (oben) und Hochpassfilter (unten) verschiedener Ordnung mit einer Grenzfrequenz von $f_1 = 500$ Hz

Im einfachsten Fall besitzt ein Filter eine feste Flankensteilheit, die üblicherweise in dB/Oktave oder dB/Dekade angegeben wird. Technisch werden Filter verschiedener Ordnung unterschieden. Je höher die Ordnungszahl n , desto höher die Flankensteilheit, die sich zu $n \cdot 20$ dB/Dekade bzw. $n \cdot 6$ dB/Oktave ergibt. Je nachdem, ob die Dämpfung mit steigender oder sinkender Frequenz zunimmt, handelt es sich um ein Tiefpass- oder ein Hochpassfilter (Abb. 13.14). Die Frequenz, an der die Dämpf-

fung 3 dB betragt, wird als Grenzfrequenz bezeichnet. Sie ist ublicherweise der einzige justierbare Parameter eines einfachen Filters.

Schaltet man beide Filtertypen hintereinander, ergibt sich in Abhangigkeit der Reihenfolge ihrer Grenzfrequenzen ein Bandpass oder eine Bandsperre (Abb. 13.15 mit entsprechenden Filtern 5. Ordnung). Zur Eliminierung einzelner Storfrequenzen (z.B. Netzbrummen) werden spezielle, schmale Bandsperren mit hoher Flankensteilheit eingesetzt, die als Notch-Filter bezeichnet werden.

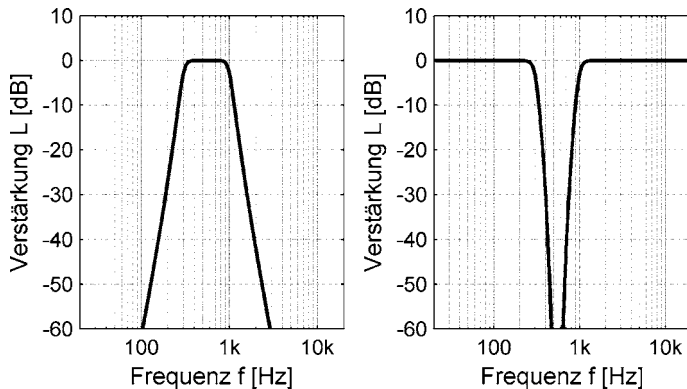


Abb. 13.15 Bandpass (links) und Bandsperre (rechts) 5. Ordnung

13.2.7.1 Equalizer

Eine in der Audiotechnik spezielle und hufig eingesetzte Art von Filtern sind Equalizer (Entzerrer). Moderne Equalizer besitzen einen Amplitudengang, der sich nach einem Abschnitt frequenzabhangiger Verstarkung bzw. Dampfung asymptotisch einem einstellbaren Zielpiegel annahert, also einen Bereich weitgehend frequenzunabhangig definierbarer Verstarkung aufweist. Im Unterschied zum einfachen Filter kann neben der Grenzfrequenz daher auch der Verstarkungspegel im entsprechenden Frequenzbereich als Parameter verandert werden (halbparametrischer Equalizer). Liegt dieser Bereich zwischen zwei definierten Grenzfrequenzen f_1 und f_2 , also inmitten des Frequenzbereichs des Nutzsignals, so spricht man in Anlehnung an die Form des Amplitudengangs von einem Glockenfilter $f_0 = \sqrt{f_1 \cdot f_2}$ (bell) oder Peakfilter, das durch eine Bandbreite $B = f_2 - f_1$, eine Mittenfrequenz und einen Verstarkungspegel L beschrieben werden kann (Abb. 13.16). Ist der Bereich nach einer Seite offen, liegt er also am linken oder rechten Rand des Ubertragungsbereichs, so wird der Equalizer als Shelf-Filter bezeichnet. Vollparametrische Equalizer ermoglichen neben der Einstellung der Mittenfrequenz auch die Einstellung des Gutefaktors Q (Abb. 13.16), wobei $Q = f_0/B$. Manche Equalizer kompensieren die bei $L \neq 0$ dB mit einer Anderung von Q einhergehende Anderung der akustischen Leistung (und damit Lautstärke) des bearbeiteten Signals. Filter mit festen Mittenfre-

quenzen und fester Güte, meistens zu Batterien mit 10 oder 30 Bändern gruppiert, werden als grafische Equalizer bezeichnet.

Da die Ausprägung des Helligkeitseindrucks, aber auch der Eindruck von klangfarblicher Inhomogenität (dröhnend, hohl), wie sie z. B. durch Resonanzeffekte hervorgerufen werden kann (Abgeschlossenheits-Dimension nach Nitsche 1978, S 27, Abb. 13.32), stark vom Betragsspektrum bestimmt wird, kann sie mit Filtern verändert werden (zu Klangfarbentheorien im Überblick vgl. Muzzulini 2006). Einfache Filter werden hauptsächlich zur Verminderung von Nebengeräuschen (z. B. Trittschall) oder unerwünschten Nutzsignalanteilen in den spektralen Außenbereichen eingesetzt und nur bedingt im kreativ-gestalterischen Sinne. Im letzteren Falle wird auf die klangfarbliche Dimension der Helligkeit Einfluss genommen. Mit parametrischen Equalizern lässt sich das Frequenzspektrum effektiv und zugleich differen-

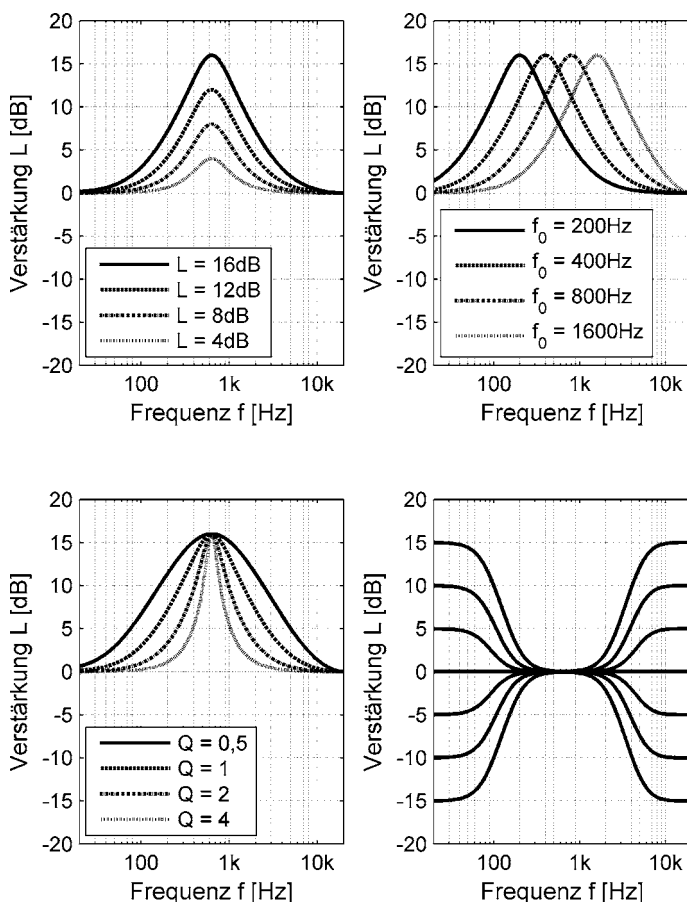


Abb. 13.16 Typen und Parameter von Equalizern: Peakfilter mit verschiedenen Verstärkungspegeln (oben links), Mittenfrequenzen (oben rechts) und Güten (unten links) sowie tiefen- und höhenwirksame Shelf-Filter mit verschiedenen Verstärkungspegeln (unten rechts)

ziert beeinflussen, entsprechend umfassend auch der Klangfarbeneindruck. In der Praxis stellen die Abhörbedingungen einen Störfaktor bei der Equalisierung dar: In einem Experiment von Börja (1978) wiesen die Differenzspektren von in verschiedenen Regieräumen angefertigten Musikmischungen Ähnlichkeiten mit der inversen kombinierten Übertragungsfunktion von Lautsprecher und Wiedergaberaum auf. Das Ergebnis einer Klangeinstellung wird außerdem durch das ursprüngliche Spektrum des zu bearbeitenden Signals beeinflusst, das wie ein Ankerreiz wirkt (Letowski 1992). Eine getrennte Einflussnahme auf Einschwingvorgänge und quasistationäre Abschnitte, bzw. generell auf Klangfarbenverläufe ist mit herkömmlichen Filtern bzw. Equalizern nicht möglich, dies kann jedoch durch digital realisierte dynamische Filter geschehen.

Einsatzgebiete des Equalizers reichen von der Entfernung bestimmter Störsignale über die Kompensation störender Formanten, das in der populären Musik verbreitete Aufhellen von Klängen und die partielle Beeinflussung der Tiefenlokalisation bis zu starken Verfremdungseffekten, etwa dem kontinuierlichen Durchstimmen schmaler Equalizer-Bänder über einen größeren Frequenzbereich (Filter-Sweep). Grafische Equalizer werden häufig für eine Frequenzganganpassung von Beschallungsanlagen an spezifische räumliche Gegebenheiten eingesetzt.

13.2.8 Verzerrer

Als Verzerrer kann man alle Audiotbearbeitungsmittel bezeichnen, die nichtlineare Verzerrungen verursachen mit dem Ziel, deutlich wahrnehmbare Artefakte zu produzieren. Es existieren vielfältige Bezeichnungen bzw. Typen wie Distortion, Over Drive, Crunch, Fuzz etc. Die nichtlinearen Verzerrungen werden in Vorverstärkern, Endstufen oder Effektgeräten durch die Übersteuerung von Verstärkerstufen verursacht, indem Röhren oder Halbleiter im nichtlinearen Bereich ihrer Kennlinie betrieben werden. Je nach Ansteuerungspegel, der mit einem meist als Gain, Boost oder Drive bezeichneten Regler einstellbar ist, wird das Signal mehr oder weniger stark deformiert bzw. abgeschnitten (clipping), was einer Pegelbegrenzung und der Addition zusätzlicher Partialtöne gleichkommt (Kap. 21.4). Bei musikalischen Tönen (mehrere Partialtöne) verursacht eine nichtlineare Verzerrung infolge von Summen- und Differenzfrequenzen Signalanteile, die gemeinsam als rau und geräuschhaft empfunden werden. Die verschiedenen Geräte oder Programme unterscheiden sich hauptsächlich in der Form der genutzten Kennlinie. Sie ist für die Härte des pegelabhängigen Effekteinsatzes und – infolge unterschiedlich starker Ausprägungen verschiedener Obertonklassen, z. B. gerad- oder ungeradzahlgiger Obertöne – die Klangfarbe des Effekts verantwortlich, die häufig zusätzlich mit Filtern verändert werden kann. Der hohe Obertongehalt und die hohe akustische Leistung infolge der Begrenzungswirkung machen mit Verzerrern bearbeitete Signale im Klangbild auffällig und durchsetzungsstark. Entsprechende Effekte werden vor allem zur Gestaltung vielfältiger E-Gitarren-Sounds eingesetzt. Neben der heute üblichen digitalen Realisierung von ursprünglich analog erzeugten Verzerrungen

etablieren sich auch typische Artefakte der digitalen Audioübertragung als Effekte. So bewirkt z. B. der Bitcrusher durch die Verringerung der Wortbreite gezielt eine Vergrößerung des Quantisierungsfehlers und damit eine klirrende, stöngeräusch-behaftete Verzerrung.

13.2.9 Enhancer

Auch Enhancer dienen der Obertonerzeugung, im Gegensatz zum Verzerrer soll die Klangbearbeitung aber nicht als deutliches Artefakt, sondern vielmehr als dezente klangfarbliche Anreicherung wahrgenommen werden. Dazu werden Obertöne aus hoch- oder bandpassgefilterten Teilen des Ausgangssignals erzeugt. Die Grenzfrequenzen der jeweiligen Filter sind mit einem meist als Tune bezeichneten Parameter wählbar, die Stärke der Obertonerzeugung mit dem Parameter Drive. Das unter Umständen stark obertonhaltige Signal wird dem Ausgangssignal nun in äußerst geringer Dosierung zugemischt (Parameter mix). Einige Prozessoren nehmen zusätzlich eine Bearbeitung auch der tiefen Frequenzen durch Kompression und/oder Verzögerung vor. Enhancer werden sowohl auf Einzelklänge als auch auf komplette Mischungen angewendet und sollen die Helligkeit und Lebhaftigkeit des Audioinhalts sowie ggf. die Sprachverständlichkeit erhöhen. Ein ähnliches Prinzip kommt in Plug-ins zur Röhren- bzw. Bandsättigungs-Simulation zur Anwendung, wobei geradzahlige bzw. ungeradzahlige Partialtöne erzeugt werden.

13.2.10 Delayeffekte

Die Verzögerung von Signalen ist digital einfach zu realisieren und Grundlage verschiedener Effektgeräte bzw. perceptiver Wirkungen. Je nach Kombination verschiedener Verzögerungszeiten, verschiedener Pegeldifferenzen, ein- und mehrmaliger Signalwiederholung, zeitlich regelmäßiger oder unregelmäßiger Signalwiederholung sowie räumlich identischer oder getrennter Wiedergabe von Original- und verzögerten Signalen kommen verschiedene psychoakustische Effekte wie Klangverfärbung, Raumeindruck, Echowahrnehmung und Lokalisationswirkungen (Summenlokalisierung, Präzedenzeffekt oder Haas-Effekt) zum Tragen. Zu Lokalisationseffekten s. Kap. 3, für durch komplexe Reflexionsmuster verursachte Raumeindrücke s. Abschn. 13.2.11.

Werden ein Signal und dasselbe um Δt verzögerte Signal summiert, resultiert eine als Kammfiltereffekt bezeichnete lineare Verzerrung, die bei den Frequenzen $f_{\text{peak}_n} = n/\Delta t$ jeweils eine Überhöhung und bei $f_{\text{dip}_n} = (2n-1)/2\Delta t$ jeweils eine Absenkung aufweist. Die relativen Pegel dieser Peaks und Dips hängen von der Pegeldifferenz zwischen Bezugssignal und verzögertem Signal ab (Abb. 13.17).

Auf perceptiver Ebene werden mit steigender Verzögerungszeit nacheinander die Wahrnehmungsqualitäten Klangfarbe, Räumlichkeit und Echo angesprochen

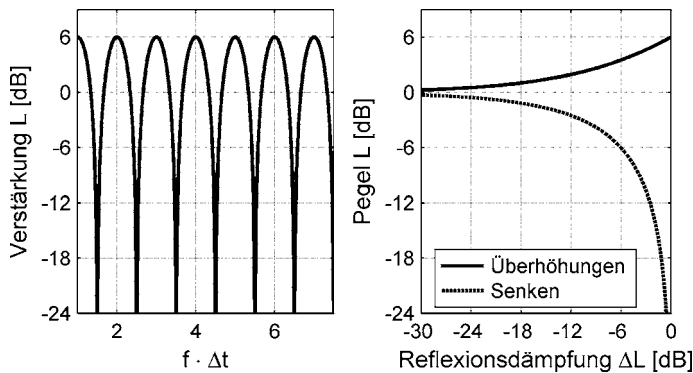


Abb. 13.17 Betragsfrequenzgang eines Kammfilters, das sich durch die Überlagerung zweier identischer zeitversetzter Signale ergibt (links), und Pegel der Überhöhungen (peaks) und Senken (dips) der Kammfilterfunktion in Abhängigkeit von der Dämpfung des verzögerten Signals (rechts)

bzw. in ihren Ausprägungen verändert. Bei breitbandigen Audioinhalten mit geringer Tonhaltigkeit kann es aufgrund der partialtonähnlichen regelmäßigen Lage der Peaks und Dips im Frequenzbereich neben dem Eindruck klangfarblicher Veränderung auch zur Tonhöhenempfindung kommen. Man unterscheidet dabei einkanalige und meist einohrige Darbietung der zeitversetzten Signale (monaural repetition pitch) und beidohrige Darbietung der interaural zeitversetzten Signale (dichotic repetition pitch); vgl. Hartmann (2005). Die genauen Verzögerungszeiten, die die Übergänge der Wahrnehmungsqualitäten markieren, und die Größe der Überlappungsbereiche hängen von vielen Faktoren ab, insbesondere von Signaleigenschaften wie Impulshaftigkeit, Tonhaltigkeit und Spektrum, der Pegeldifferenz zwischen Bezugssignal und verzögertem Signal und der Schulung des Gehörs (vgl. Kap. 3.3.2).

Kammfiltereffekte können auch bei Mikrofonaufnahmen durch Überlagerung des Direktsignals mit akustischen Reflexionen, z. B. durch ein Sprecherpult, verursacht sein. Brunner et al. (2007) zeigten, dass **Expertenhörer bei Kammfiltereffekten** durchschnittlich noch Pegeldifferenzen von 18 dB zuverlässig entdecken können, im Einzelfall sogar Pegeldifferenzen von 27 dB. Die Versuchspersonen reagierten bei Verzögerungszeiten zwischen 0,5 und 3 ms am empfindlichsten. Dies entspricht einer Schallweglänge von 0,17 bis 1 m, die bei der Mischung von Mikrofonsignalen vermieden werden sollte.

Verzögerungsgeräte realisieren den als Parameter Delay Time wählbaren Zeitversatz heute durch eine digitale Zwischenspeicherung des Signals. Die Verzögerungszeit kann um ihren justierten Wert herum variiert werden (modulation), was technisch einer geänderten Auslesegeschwindigkeit und perzeptiv einer Tonhöhenänderung gleichkommt. Die Modulation kann entweder periodisch durch einen Tieffrequenzoszillator (Parameter speed), signalverlaufsabhängig durch einen Hüllkurvengenerator (envelope) oder zufällig (random) gesteuert werden. Die Modulationstiefe ist mit dem Parameter Depth einstellbar. Das verzögerte und ggf. frequenzmodulierte Signal kann nun dem Originalsignal dosiert zugemischt werden

(mix) und zugleich zum Eingang der Verzögerungsstufe rückgekoppelt werden (feedback). Aufwändige Geräte beinhalten mehrere parallele Verzögerungsstufen oder ermöglichen zusätzlich Dynamik-, Filter- und Panorama-Modulationen.

Tabelle 13.2 Typische Delayeffekte

Effekt / Artefakt	delay time	modulation	feedback
Verzögerung single delay	> 40 ms	nein	nein
Verdopplung double tracking	20–40 ms	nein	nein
Echo	> 100 ms	nein	ja
multitap delay	mehrere diskret justierbare Verzögerungszeiten	beliebig	beliebig
Chorus	15–30 ms	ja	nein
Flanger	1–10 ms	ja	ja

Durch die Wahl bestimmter Kombinationen von Parametern lassen sich verschiedene klassische Delay- oder Modulationseffekte erzeugen (Tabelle 13.2). Eine einfache Verzögerung (single delay) mit einer Delayzeit bis zu einigen Sekunden vermittelt den Eindruck eines Rückwurfs, wie er ähnlich in der Natur an großen und mehr oder weniger weit entfernten reflektierenden Flächen auftritt (Stadion, Waldrand, Felswand); Zeiten im Bereich der Echschwelle werden häufig zur Verdopplung von Gesangsstimmen gewählt (double tracking). Delayzeiten zwischen 100 ms und einigen Sekunden ergeben in Verbindung mit leichtem oder starkem Feedback den klassischen Echo-Effekt. Mehrere unabhängig voneinander in Delay time und Mix einstellbare Verzögerungseinheiten ermöglichen die Generierung auch nicht-periodischer Wiederholungsmuster bzw. Rhythmen (multitap delay). Chorus- und Flanger-Effekt basieren auf der Mischung von Original- und verzögertem Signal, wobei die Verzögerungszeit periodisch variiert wird, was auf eine waagerechte Streckung und Stauchung der Kammfilterkurve in Abb. 13.17 hinausläuft. Die beiden Effekte unterscheiden sich in der Wahl der Delayzeit und im Feedback-Einsatz. Beim Flanger-Effekt dominiert perzeptiv die veränderliche komplexe Klangverfärbung mit Tonhöhencharakter, der Chorus hingegen schafft neben einer dunkleren, weniger auffälligen Klangfärbung einen Eindruck von Räumlichkeit und verleiht dem bearbeiteten Signal so mehr Klangfülle. Die Effektsignale müssen nicht mit dem Originalsignal gemischt, sondern können auch auf anderen Kanälen ausgespielt werden, um stereofon eingesetzt werden zu können, was z. B. typisch für das Double Tracking ist. Delayeffekte werden hauptsächlich in der Produktion von Populärmusik eingesetzt und sind häufig musikstrukturell abgestimmt. Die Delayzeit, die einem bestimmten Notenwert bei einem bestimmten Tempo entspricht, berechnet sich einfach durch $T=k \cdot 60/b$ [s] aus dem Tempo b in bpm (beats per minute) und der Notenwert-Konstante k in Bruchteilen des Grunds Schlag-Wertes, meist der Vier-

telnote (z. B. ♩ = 0,25 bzw. ♩ = 2). Bei der Übertragung klassischer Musik werden Delays auch für die Verzögerung der Signale von Stützmikrofonen eingesetzt (Kap. 10.5.3).

13.2.11 Nachhall

Geräte zur Erzeugung von künstlichem Nachhall erfüllen eine Vielzahl von Funktionen. Bei der Produktion klassischer Musik werden sie in der Regel zur Verlängerung der Nachhalldauer eingesetzt, wenn die Aufnahme in einer akustisch nicht vollständig befriedigenden Umgebung stattfindet. In der Populärmusik, wo meist mit dicht mikrofonierten Signalen mit geringem räumlichen Anteil gearbeitet wird, entsteht die räumliche Perspektive (Entfernung, Raumgröße, Halligkeit) erst durch die Mischung des aufgenommenen Signals mit künstlichem Nachhall, häufig in Kombination mit einzelnen Reflexionen, wie sie durch separate Delayeffekte (Abschn. 13.2.10) oder durch den Nachhallalgorithmus selbst generiert werden. Auch wenn künstlicher Nachhall heute praktisch ausschließlich durch digitale Effektgeräte erzeugt wird, ist ein kurzer Rückgriff auf die Vorgeschichte analoger Verfahren allein deshalb interessant, weil einige dieser Verfahren (wie die Hallplatte) mit der ihnen eigenen Klangcharakteristik auch von digitalen Algorithmen simuliert und eingesetzt werden.

Das älteste Werkzeug zur Erzeugung von zusätzlichem Nachhall waren Hallräume, wie sie unmittelbar nach Ablösung des mechanisch-akustischen Aufnahmeverfahrens durch eine elektroakustische Übertragungskette um 1925 eingerichtet wurden. Die Möglichkeit der elektrischen Mischung von an verschiedenen Orten aufgenommenen, mikrofonierten Klängen wurde genutzt, um Signale über Lautsprecher in einen Raum mit stark reflektierenden Wänden einzuspielen, über Mikrofone zurückzuführen und dem Quellsignal als räumlichen Anteil zuzumischen. In den 1930er Jahren war dies gängige Praxis in den Aufnahmestudios bei Rundfunk und Schallplatte, ebenso wie beim Film (Rettinger 1945).

Nach Einführung der magnetischen Aufzeichnungstechnik wurden in den 1950er Jahren eine Reihe von Nachhallerzeugern entwickelt, die mit magnetisch beschichteten Rädern arbeiteten, deren Oberfläche mit einer Reihe von Aufnahme- und Wiedergabeköpfen bespielt und ausgelesen wurde und so eine abklingende Folge verzögerter Reflexionen produzierte, deren zeitliche Struktur durch die Position der Tonköpfe vorgegeben war. Solche Apparate kamen sowohl im Studiobetrieb zum Einsatz (Axon et al. 1957) als auch zur Nachhallverlängerung von Konzertsälen (Kap. 5.5, Abb. 5.46).

Universelle Verbreitung fand die 1957 von der deutschen Firma EMT eingeführte Hallplatte. Eine rechteckige, an ihren Eckpunkten eingespannte Stahlplatte von 1x2 m Kantenlänge wurde durch einen elektrodynamischen Wandler zu Biegeschwingungen angeregt, die sich entlang der Plattenoberfläche ausbreiteten und an den Kanten reflektiert wurden. Dieses Reflexionsmuster wurde an einer anderen

Stelle der Platte piezoelektrisch abgenommen und in ein elektrisches Signal zurückgewandelt. Durch eine an die Hallplatte angenäherte Dämmplatte konnte die Nachhallzeit zwischen 1 und 5 s eingestellt werden (Kuhl 1958). Die Hallplatte EMT 140 lieferte zunächst nur ein monofones Signal, seit 1961 war sie mit einem stereofonen, seit 1973 auch mit einem quadrofönen Abnehmer erhältlich (EMT 1983). Das gleiche Prinzip der Nachhallerzeugung kam auch beim Nachfolgemodell, der 1971 eingeführten Hallfolie EMT 240 zu Einsatz. Dabei wurde eine nur noch 27x29 cm große Folie aus einer Gold-Legierung piezoelektrisch zu Schwingungen angeregt und deren Reflexionsmuster durch einen dynamischen Wandler abgetastet. Der Vorteil gegenüber der Hallplatte lag neben der Kompaktheit des Systems vor allem in der höheren, einem natürlichen Raum ähnlicheren Eigenfrequenzdichte der Folie. Die geringe Eigenfrequenzdichte ist hauptverantwortlich für den metallischen und tendenziell kleinräumigen Klang der Hallplatte.

Algorithmen zur Erzeugung von Nachhall durch digitale Signalverarbeitung wurden bereits Anfang der 1960er Jahre vorgeschlagen (Schroeder 1961). Der sog. Schroeder-Algorithmus (Schroeder 1962, Abb. 13.18) verwendete vier parallel geschaltete Rückkopplungsschleifen (IIR-Filter, Kap. 15.2.2), die aufgrund ihres Frequenzgangs auch als Kammfilter (comb filter) bezeichnet werden. Durch geschickte Wahl der Verzögerungen m_1 bis m_4 kann das Filter so entworfen werden, dass sich die Berge und Täler des Frequenzgangs gegenseitig in etwa kompensieren. Durch zwei nachgeschaltete Allpass-Filter mit linearem Frequenzgang (Kap. 15.2.2, Abb. 15.18) wird die Anzahl der Reflexionen jeweils noch einmal verdreifacht.

Anfang der 1960er Jahre stand noch keine Rechnerarchitektur zur Verfügung, um Algorithmen dieses Typs in Echtzeit auszuführen. In der Computermusik, für die nicht notwendigerweise in Echtzeit realisierte *Berechnung* von Kompositionen, fanden rekursive Nachhallalgorithmen jedoch sehr früh Eingang, etwa in das Mitte

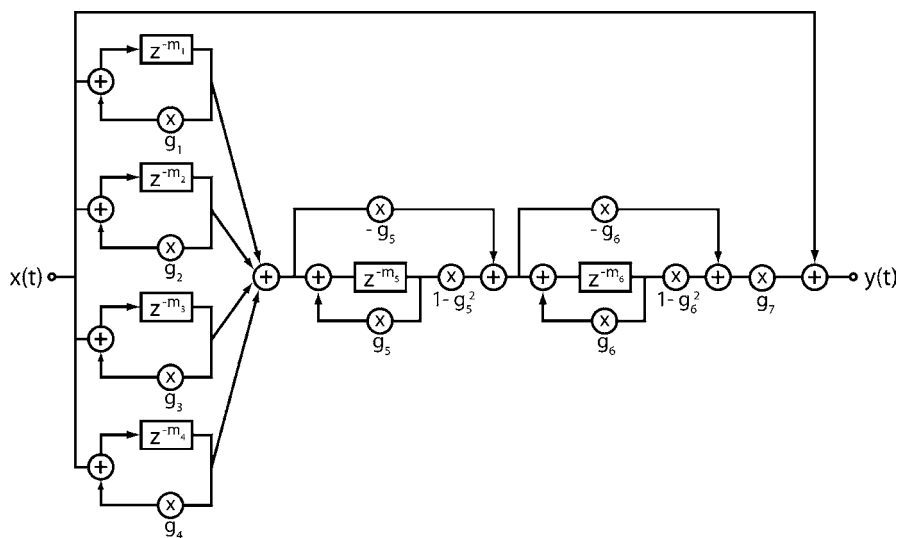


Abb. 13.18 Schroeder-Algorithmus zur Nachhallsynthese

der 1960er Jahre am Computer Center for Research in Music and Acoustics (CCRMA) in Stanford von John Chowning entwickelte Programm zur räumlichen Steuerung von Schallquellen (Chowning 1971), das zum ersten Mal bei der Komposition *Turenas* (1972) zum Einsatz kam. Mitte der 1970er Jahre wurde die Erzeugung von digitalem Nachhall in Echtzeit auf einem industriellen Großrechner demonstriert (Baeder u. Blesser 1975), bevor wiederum der Firma EMT mit dem 1976 vorgestellten digitalen Nachhallgerät EMT 250 die erste Implementierung eines rekursiven Nachhallalgorithmus auf einem tonstudiotauglichen Digitalrechner gelang. Es erlaubte einen viel weitergehenden Zugriff auf einzelne Parameter und Klangeigenschaften des Nachhalls als dies bei analogen Geräten möglich war. In der Folgezeit stieg mit der Rechenleistung von Mikroprozessoren sowohl die Übertragungsqualität der Nachhallerzeuger (das EMT 250 arbeitete mit 24 kHz Abtastfrequenz und 12 bit Wortbreite) als auch die Komplexität der implementierten Algorithmen. So erweiterte (Moorer 1979) den Schroeder-Algorithmus um zwei weitere, parallel geschaltete Comb-Filter und einen Tiefpass erster Ordnung in jedem Rückkopplungszweig zur Bedämpfung hoher Frequenzen, wie sie in der Realität durch die Luftabsorption und das Absorptionsverhalten der Wände gegeben ist. Außerdem werden Flatterechos, wie sie bei kurzen transienten Signale hörbar werden können, in ihrer Wirkung abgemildert.

Das von Jot u. Chaigne (1991) analysierte General Feedback Delay Network kann als Verallgemeinerung der Algorithmen von Schroeder und Moorer verstanden werden. Es erreicht eine maximale Dichte und eine maximale Komplexität in der zeitlichen Struktur der Reflexionsfolge, indem die Ausgänge parallel geschalteter Delaylines auf alle Eingänge rückgekoppelt und gemischt werden. Die Gewichte der Rekursion werden durch eine Feedback Matrix kontrolliert. Durch eine geeignete Wahl der Koeffizienten für das Feedback, für ein in jeder Rückkopplungsschleife befindliches Absorptionsfilter und für ein am Ausgang liegendes Korrekturfilter kann, im Gegensatz zum Schroeder-Algorithmus, das zeitliche und spektrale Verhalten des Nachhalls unabhängig voneinander konfiguriert werden.

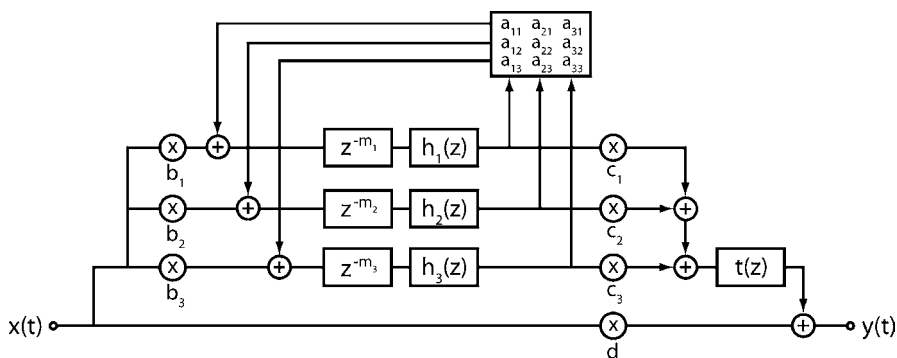


Abb. 13.19 General Feedback Delay Network zur Erzeugung von nachhallähnlichen Reflexionsmustern mit einer durch die Feedback Matrix $A = (a_{ij})$ gewichteten Parallelschaltung von Rückkopplungsschleifen, Absorptionsfiltern $h_i(z)$ zur Simulation der Absorption von Luft und Wänden und einem Korrekturfilter $t(z)$, nach Jot u. Chaigne (1991)

Auch wenn das genaue Design moderner Nachhallalgorithmen bei allen Herstellern ein mehr oder weniger streng gehütetes Firmengeheimnis ist, darf man davon ausgehen, dass sie im Prinzip auf einem in Abb. 13.19 skizzierten Algorithmus beruhen. Komplexe Implementierungen wie die Programme der Fa. Lexicon, die sich seit Ende der 1970er Jahre mit den Modellen 224L (1978), 480L (1986) und 960L (2000) als Marktführer etabliert hat, verwenden auch zeitvariante Elemente wie eine zufallsgesteuerte zeitliche und spektrale Modulation früher Reflexionen, wie sie in realen Räumen durch eine Bewegung der Schallquelle oder des Hörers und durch die frequenzabhängige Richtwirkung der Quellen entstehen.

Geräte zur Nachhallsynthese durch IIR-Filter bieten dem Benutzer eine große Palette an Parametern zur Konfiguration des erzeugten Reflexionsmusters. Dazu gehört zunächst eine Auswahl verschiedener Raumtypen (church, hall, chamber, room), durch die Grundparameter des Nachhallprogramms wie Nachhallzeit, die Struktur früher Reflexionen sowie Reflexionsdichte und Eigenfrequenzdichte vor-eingestellt werden. Auch eine Simulation des Verhaltens der Nachhallplatte (plate) gehört zum Repertoire fast aller Nachhallprogramme. Jedes dieser Programme bietet einen Parameter für die Raumgröße (size), der häufig als Master-Regler eine in sich konsistente Einstellung anderer Parameter bewirkt, für die Nachhallzeit (reverb time), einen Faktor für die Änderung der Nachhallzeit bei tiefen Frequenzen (bass multiply), eine Grenzfrequenz für die Dämpfung hoher Frequenzen (rolloff) und die Abnahme der Nachhallzeit zu hohen Frequenzen (treble decay), Parameter zur Steuerung der Hüllkurve im Ein- und Ausschwingvorgang (shape, spread), eine Zeitverzögerung des Nachhalls gegenüber dem Eingangssignal (predelay) und Parameter zur Erzeugung einzelner früher Reflexionen mit definierter Zeitverzögerung und definiertem Pegel.

Seit Ende der 1990er Jahre ist die Rechenleistung von Prozessoren hoch genug, um künstlichen Nachhall nicht nur durch rekursive Algorithmen zu erzeugen, sondern auch sog. Faltungsalgorithmen in Echtzeit auszuführen. Faltungsalgorithmen sind FIR-Filter, deren Filterkoeffizienten durch die Abtastwerte der Impulsantwort von realen Räumen gebildet werden. Durch diesen Prozess, der um ein Vielfaches höhere Anforderungen an die Rechenleistung stellt als rekursive Algorithmen (bei einer Nachhallzeit von 2 s und 48 kHz Abtastfrequenz sind bereits 96.000 Filterkoeffizienten erforderlich), kann jedem Eingangssignal ein Nachhallverlauf aufgeprägt werden, der sich zeitlich und spektral nicht von einem im realen Raum aufgenommenen Nachhall unterscheidet. Seit der Einführung einer algorithmischen Lösung, die gleichzeitig den Aufwand und die Latenz der Berechnung minimiert, d.h. das Ergebnis der Faltung am Ausgang des Prozessors praktisch verzögerungsfrei zur Verfügung stellt (Gardner 1995), wurde auch dieses Verfahren zunächst auf externen DSP-Architekturen realisiert (Sony DRE-777, Yamaha SREV1). Seit einigen Jahren ist es als Funktion digitaler Audioworkstations oder in Form von Plug-ins auch auf Host-Prozessoren lauffähig und wird meist mit einer Bibliothek von Impulsantworten ausgeliefert, die ein breites Spektrum raumakustischer Umgebungen abdecken.

Im Gegensatz zu rekursiven Nachhallalgorithmen bietet ein Faltungshall allerdings nur begrenzten Zugriff auf die Struktur des Nachhalls. In der Regel ist nur ein Predelay verfügbar, um den Nachhall gegenüber dem Direktsignal zu verzögern,

sowie ein Parameter zur Modifikation der Nachhallzeit, der das Ausklingen der Impulsantwort verkürzt oder verlängert.

Im Hinblick auf eine Simulation des Reflexionsverhaltens natürlicher Räume kann die Nachhallsynthese mit Einführung dieser Faltungsalgorithmen im Grundsatz als gelöst angesehen werden, auch wenn zeitvariante Elemente, wie sie etwa durch die Bewegung von Quelle und Hörer entstehen, durch eine Impulsantwort, die nur eine Momentaufnahme des Übertragungssystems darstellt, nicht abgebildet werden. Weil zudem häufig nicht die Simulation eines natürlichen Raums im Vordergrund steht, sondern das Hinzufügen spezifischer räumlicher Anteile zu einem bereits raumbehafteten Signal, sind auch rekursive Nachhallalgorithmen nach wie vor ein unverzichtbares Werkzeug der Klanggestaltung, da sie bei der räumlichen Effektbearbeitung einen größeren Gestaltungsspielraum ermöglichen.

13.2.12 Pitch Shifting

Die Funktion von Pitch-Shifting-Effekten besteht gemäß ihrer Bezeichnung in der Verschiebung der Tonhöhe von Audiosignalen. Die entsprechende Signalverarbeitung beruht auf dem Prinzip des Phase-Vocoders, eines FFT-basierten Ansatzes (Bernsee 2005), oder der Granularsynthese, die mit kleinen Abschnitten des digitalen Audiomaterials (grains) im Zeitbereich operiert (Roads 2001). Man unterscheidet zwischen einfachen Pitch Shiftern, bei denen Tonhöhenverschiebungen fest bzw. für verschiedene Zeitabschnitte voreinstellbar sind, Harmonizern®, die durch Rückkopplung Mehrklänge aus geschichteten identischen Intervallen erzeugen können, und Geräten zur automatischen Tonhöhenkorrektur, die durch eine dynamische Steuerung der Parameter in Abhängigkeit von den Ergebnissen einer Tonhöhenbestimmung geschieht (vgl. z. B. Antares Audio Technologies 2006). Zu den justierbaren Parametern gehören die Tonhöhendifferenz (pitch, detune) bzw. die Stärke einer automatischen Korrektur (amount), die Korrekturgeschwindigkeit (attack, retune speed), die Tonhöhenunterschiedsschwelle, unter der eine Korrektur stattfindet (window), sowie Typ und Stimmung der Bezugstonart bzw. -skala (scale). Darüber hinaus bieten heutige Geräte ausführliche Analysefunktionen und -darstellungen, erweiterte Tonhöhenfunktionen (z. B. Vibratogestaltung) und integrierte Dynamikfunktionen (z. B. De-Esser). Werden Tonhöhenverschiebungen von wenigen Cents gewählt, ergibt sich durch Mischung mit dem unbearbeiteten Signal aufgrund der entstehenden Schwebung ein chorus-ähnlicher Effekt. Durch Tonhöhenverschiebungen in der Größenordnung musikalischer Intervalle wird die entsprechende Stimme transponiert und so in das musikalisch-strukturelle Gefüge eingegriffen. Durch mehrfache simultane Tonhöhenverschiebung können aus einer Stimme Mehrklänge gebildet werden. Speziell auf die Generierung von Ober- und/oder Unteroktaven ausgelegte Geräte werden als Octaver bezeichnet. Viele Geräte ermöglichen statt einer intervallstarrten Transposition die Berücksichtigung von Tonarten oder in Einzeltönen vorgegebenen Skalen (intelligent pitch shift). Ein perzeptiv meist unerwünschter Nebeneffekt deutlicher Tonhöhenveränderungen ist die

gleichzeitige Verschiebung von Formanten, was z. B. den bekannten Mickey-Mouse-Effekt bewirkt. Moderne Algorithmen ermöglichen durch die Anwendung von Frequenzbereichsverfahren (vgl. 15.3.5) die unabhängige Verschiebung von Tonhöhe und Formanten bzw. die Vermeidung von Formantverschiebungen (vgl. etwa Hoenig und Neubäcker 2006). Auch ein solches Processing produziert noch hörbare Artefakte (insbesondere bei automatisch korrigierten gebundenen Tonhöhenwechseln). Mittlerweile sind diese allerdings klangästhetisch etabliert und können als eigene Kategorie von Effekten angesehen werden („Cher sound“).

13.2.13 Time Stretching

Time-Stretching-Algorithmen ermöglichen die Veränderung der Länge eines Audio-signals bzw. seines Abspieltempos ohne Tonhöhenänderung und basieren auf denselben Signalverarbeitungsprinzipien wie das Pitch Shifting (Abschn. 13.2.12, Kap. 15.3.5). Time-Stretching-Algorithmen sind sowohl als eigenständige Anwendungen verfügbar als auch Bestandteil der gängigen integrierten Audibearbeitungssysteme. Wählbar sind der Zeitskalierungsfaktor und ggf. für verschiedene Audioinhalte optimierte Analysemodi zur Minimierung von Artefakten. Erst der kombinierte Einsatz von Time Stretching, Pitch Shifting und Schnitt erweitert die Möglichkeiten des Samplings bzw. der nonlinearen Audioproduktionsweise beträchtlich: Audiomaterial kann so nicht nur klangfarblich und dynamisch in Medienproduktionen eingepasst werden, sondern auch hinsichtlich Zeitpunkt, Wiedergabeausschnitt, Spieldauer und Tonhöhe.

13.2.14 Phaser

Der Phaser ist ein Modulationseffekt, der im klanglichen Ergebnis dem Flanger ähnelt, da er ebenfalls auf der Überlagerung von Original- und zeitverschobenem Signal beruht. Allerdings wird beim Phaser nicht der Zeitversatz, sondern die Phasenlage der Frequenzkomponenten moduliert, was auf ein nicht mehr gleichabständiges Auftreten der Peaks und Dips der Kammfilterfunktion hinausläuft (vgl. Abb. 15.30 rechts). Technisch wird der Phaser durch eine Serie von Allpass-Filtern realisiert (Hartmann 1978, Abschn. 15.3.3). Neben der stets einstellbaren Modulationsgeschwindigkeit (rate, speed) sind je nach Bauart auch die Parameter Modulationstiefe (depth, intensity, amount), Modulationsmittelpunkt (manual, sweep) und Rückkopplung (feedback, regen, resonance) justierbar, sowie ggf. eine Phaseninvertierung des verschobenen Signals (phase, mode, colour) und die Anzahl der Filterstufen (stage) schaltbar. Für eine Veranschaulichung wesentlicher Parameter vgl. z. B. Moog Music (2003a). Der in den 1960er und 70er Jahren beliebte Effekt wird häufig für die Bearbeitung von Gitarren- und Synthesizersounds eingesetzt und ist daher oft in Bodeneffektgeräten anzutreffen.

13.2.15 Ringmodulator

Ein Ringmodulator multipliziert zwei Wechselspannungen und erzeugt auf diese Weise deren Differenz- und Summenfrequenz (Zweiseitenbandmodulation mit unterdrücktem Träger). Der Name Ringmodulator geht auf seine analoge Konstruktion durch vier gleichsinnig im Ring angeordnete Dioden zurück. Eine vereinfachte, zur Verarbeitung von Rechtecksignalen geeignete digitale Realisation läuft auf eine logische Exklusiv-Oder-Verknüpfung der Signale hinaus. Nachdem ein als Audio-Effektgerät ausgelegter Ringmodulator in der Regel nur ein Signal verarbeiten soll, enthält er meistens einen lokalen Oszillator zur Erzeugung des Trägersignals (vgl. z. B. Moog Music 2003b). Dessen mittlere Frequenz ist in einem weiten Bereich (z. B. 0,1 Hz bis 5 kHz) einstellbar (frequency) und oft zusätzlich durch einen Tief-frequenz-Oszillator (LFO) modulierbar, der ggf. verschiedene Wellenformen erzeugen kann (waveform). Die LFO-Parameter sind Rate oder Speed für die Modulationsfrequenz und Amount, Intensity oder Depth für die Modulationstiefe der Trägerfrequenz. Die Stärke der eigentlichen Ringmodulation durch das eingespeiste Audiosignal wird mit dem Parameter Drive oder Level und die Mischung von ring-moduliertem und originalem Signal mit dem Parameter Mix eingestellt. Manche Geräte beinhalten zusätzlich Filter zur Klangfarbenbeeinflussung. Durch die produ-

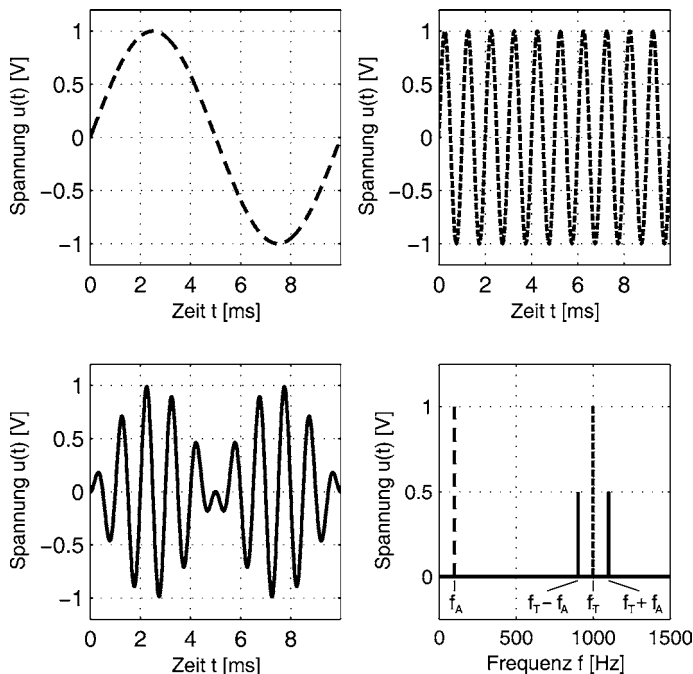


Abb. 13.20 Wirkung der Ringmodulation zweier Sinussignale mit den Frequenzen $f_A = 100$ Hz (oben links) und $f_T = 1000$ Hz (oben rechts) im Zeitbereich (unten links) und Frequenzbereich (unten rechts). Die Eingangsfrequenzen (gestrichelt) sind im modulierten Signal nicht mehr enthalten.

zierten Summen- und Differenztöne führt eine kontinuierliche Frequenzveränderung des Eingangs- oder Trägersignals zu einem Klangeffekt, wie er ähnlich beim Durchstimmen eines Kurzwellenempfängers auftritt. Bei statischen Eingangsfrequenzen ähneln die durch die Ringmodulation erzeugten nichtharmonischen Frequenzspektren (vgl. Abb. 13.20) perzeptiv denen von Idiophonen. Je nach Einstellung lässt sich insoweit eine breite Palette von Klangverfremdungen erzeugen, die von dezenten Vibrato- und Verstimmungseffekten über glockenähnliche Klänge bis zur stark geräuschhaften, völligen Zerstörung der ursprünglichen Klangstruktur reichen. Da periodische bzw. tonhaltige Signale, wie sie Oszillatoren erzeugen, einem solchen destruktiven Effekt der Ringmodulation weniger schnell unterliegen als geräuschhafte, werden Ringmodulatoren vorzugsweise für die Bearbeitung von Synthesizer-Sounds eingesetzt.

13.2.16 Tremolo und Vibrato

Der Tremolo- und der Vibrato-Effekt sind eine modellhafte Nachbildung der gleichnamigen Vorgänge beim Instrumentalspiel bzw. Gesang. Der Begriff Tremolo bezeichnet spieltechnisch eine periodische Tonrepetition (z.B. bei der Gitarre) oder Lautstärkeänderung (z.B. beim Akkordeon), der Begriff Vibrato eine periodische Tonhöhenänderung (z.B. bei Streichinstrumenten oder Gesang), die sich auch klangfarblich auswirkt (Meyer 1991). Spieltechnisch sind beide Vorgänge in der Regel nicht völlig getrennt voneinander umsetzbar. In der Audibearbeitung wird ein Tremolo erzeugt, indem das Audiosignal durch einen Tieffrequenzoszillator (LFO) mit sinus- oder dreieckförmiger Wellenform amplitudenmoduliert wird (typischerweise mit einer Frequenz von 3-10 Hz), was analog durch einen spannungsgesteuerten Verstärker (VCA) und digital durch eine Multiplikation der Amplitudenwerte von Audio- und Modulationssignal erreicht wird (Abschn. 15.3.2). Zur Erzeugung eines Vibratos wird statt der Amplitude die Frequenz des Audiosignals moduliert, im Falle analoger Tonerzeugung durch die Verwendung eines spannungsgesteuerten Oszillators (VCO), im Falle beliebiger Audiosignale digital durch die Änderung der Auslesegeschwindigkeit der Audiodaten, wie es bei einigen Delayeffekten geschieht (Abschn. 13.2.10 und 15.3.2). Für beide Effekte lassen sich die Modulationsfrequenz (rate oder speed) und die Modulationstiefe (depth oder intensity) einstellen.

13.2.17 Leslie-Kabinett (rotary speaker)

Die von Donald James Leslie entwickelten und seit den 1940er Jahren eingesetzten Zwei-Wege-Lautsprecher mit rotierenden Schalltrichtern (Abb. 13.21) verursachen durch die periodisch variierende Entfernung der Schallquelle zum Hörer und zu den reflektierenden Flächen des Wiedergaberaums zugleich Änderungen der Tonhöhe

(aufgrund von Dopplereffekten), der Lautstärke, der Klangfarbe, des Raumeindrucks und geringfügig der Lokalisation. Durch diese Variation vieler wesentlicher Klangmerkmale gewinnt das übertragene Signal an Komplexität. Vor allem statische Klänge lassen sich so bereichern, weswegen Leslie-Kabinette vor allem für elektrische Orgeln (typisch: Hammond-Orgel) eingesetzt werden.

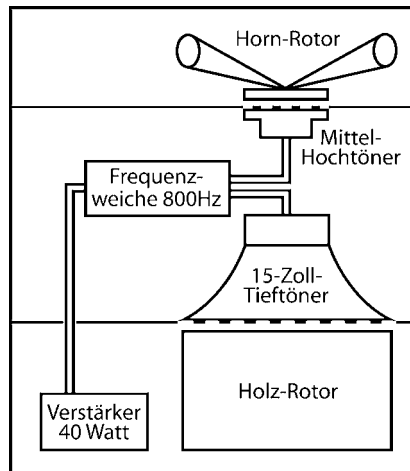


Abb. 13.21 Prinzipschaltbild des Leslie-Modells 122 mit Verstärkereinheit, Frequenzweiche, Lautsprechern und Rotoren

Es existieren zahlreiche Modelle (vgl. die Dokumentation von Mikael 2005). Schalter bzw. Regler ermöglichen die Zuschaltung des Geräts, die Lautstärkeeinstellung, die Einschaltung der Rotation der Schalltrichter (die mit verschiedenen Geschwindigkeiten erfolgt) sowie bei vielen Modellen die Wahl der Rotationsgeschwindigkeit in zwei Stufen (chorale/tremolo). Elektronische Nachbildungen des Leslie-Kabinetts nutzen v.a. den Doppler- und den Tremolo-Effekt aus. Sie bieten im Hinblick auf Größe und Gewicht eines originalen Geräts Vorteile, allerdings geht die hier besonders wichtige Wirkung allseitiger Schallabstrahlung verloren, die nur in der Live-Darbietung zur Geltung kommen kann. Eine dem Leslie-Kabinett verwandte, auf die Übertragung von Gitarrenklängen ausgerichtete und mit nur einem Rotor arbeitende Konstruktion ist das Fender Vibratone.

13.2.18 Wah-Wah

Das Wah-Wah ist ein Bandpassfilter hoher Güte (Resonanzfilter) mit variabler Mittenfrequenz. Damit diese während des Instrumentalspiels verändert werden kann, ist sie zumeist durch ein Fußpedal steuerbar, das mechanisch an ein Potentiometer, einen Spulenkern oder eine Blende vor einem Fotowiderstand gekoppelt ist. Der Bereich der Mittenfrequenz umfasst etwa 250 Hz bis 2 kHz. Je nach Gerätetyp ist die Einstel-

lung des Regelbereichs, der Filtergüte und des Mischungsverhältnisses zwischen Original- und Effektsignal möglich. Während der Ruhestellung des Pedals wird das Eingangssignal durchgeschliffen (bypass). Verschiebt man die Mittenfrequenz nach oben und unten, ergibt sich bei breitbandigen Signalen ein Klangeffekt, der durch die englische Aussprache des Wortes Wah-Wah anschaulich beschrieben ist. Die klangliche Verwandtschaft mit der menschlichen Stimme erklärt sich aus dem ähnlichen Frequenzbereich von Effekt-Mittenfrequenzen und ersten Formanten der Vokale. Einige Geräte ermöglichen eine Steuerung der Mittenfrequenz durch die Zeithüllkurve des zu bearbeitenden Audiosignals (Auto-Wah) oder eines anderen Steuersignals, was z. B. eine genaue Rhythmisierung des Effekts ermöglicht. Das Wah-Wah ist zwar seit den 1960er Jahren ein typischer Gitarreneffekt, geht aber auf eine im Jazz praktizierte Spieltechnik für Blechblasinstrumente zurück, bei der ein spezieller Wah-Wah-Dämpfer während des Spiels mit der Hand auf- und abgedeckt wird (vgl. Bertsch 1994) und die schon in den 1930er Jahren populär wurde.

13.2.19 Vocoder

Unter dem Einfluss des Berliner Akustikers Karl Willy Wagner wurde in den 1930er Jahren an den amerikanischen Bell Telephone Laboratories der Vocoder entwickelt und 1939 von Homer Dudley vorgestellt. Er diente ursprünglich dazu, Sprache kodiert zu übertragen, wurde aber schon in den 1940er Jahren für die Erzeugung von Klangeffekten in Hörspielen und im Theater eingesetzt (vgl. Ungeheuer und Supper 1995). Am Quelle-Filter-Modell für Sprache orientiert, nimmt ein Vocoder nacheinander eine Spektralanalyse und eine Spektralsynthese vor (Abb. 13.22).

Im Analyseteil wird das eingespeiste Nutzsignal (program input, analysis input, speech input), typischerweise Sprache, nach optionaler Dynamikkompression durch eine Filterbank in (meist 8 bis 32) Frequenzbänder aufgeteilt. Aus den gefilterten Audiosignalen werden durch Envelope-Follower, bestehend aus Gleichrichter und Tiefpass, Zeithüllkurvensignale erzeugt – Steuerspannungen, die ursprünglich für die Übertragung bestimmt waren. Zusätzlich wird die Sprachgrundfrequenz ermittelt sowie in einem Voiced-Unvoiced-Detector (VUD) bestimmt, ob ein stimmhafter oder ein stimmloser Sprachlaut vorliegt. Einige Modelle enthalten zusätzlich eine Sprachpausenerkennung (speech detector, nicht abgebildet). Die Analyseergebnisse werden ebenfalls als Steuerspannungen übertragen. Die Detektion der Stimmhaftigkeit erfolgt meist durch einen Pegelvergleich zweier Frequenzbänder ($f_1 > 5$ kHz, $f_2 < 1$ kHz), für die Grundfrequenzbestimmung werden unterschiedliche Verfahren angewandt.

Im Syntheseteil wird zunächst ein Ersatzsignal generiert. Zur Nachahmung stimmhafter Laute wird hierfür von einem Tongenerator (VCO) ein Rechteck- oder zumeist Sägezahnsignal erzeugt, dessen Frequenz der übertragenen Steuerspannung folgt oder alternativ über ein Keyboard spielbar ist; zur Erzeugung der stimmlosen, rein geräuschhaften Anteile wird auf einen Rauschgenerator umgeschaltet. Für die Anwendung des Vocoders als Audioeffekt ist jedoch die Einspeisung eines beliebigen externen Audiosignals (carrier input, synthesis input, replacement signal) ty-

pisch, das alternativ als tonhaltige Komponente des Ersatzsignals fungiert. Bei einigen Geräten folgt der Pegel des Rauschgenerators dem des externen tonhaltigen Signals, oder es besteht zusätzlich die Möglichkeit der alternativen Einspeisung eines geräuschhaften Signals (nicht abgebildet). Das Ersatzsignal wird durch Filter in spektrale Komponenten zerlegt. Jede dieser Komponenten wird mithilfe eines spannungsgesteuerten Verstärkers (VCA) von einem über die Matrix zugeordneten Zeithüllkurvensignal der Analysestufe amplitudenmoduliert. Die Summation der modulierten Komponenten ergibt das spektral modulierte Ersatzsignal (vocoder output), dem als Perzept die zeitlichen Klangfarbenverläufe des Nutzsignals aufgeprägt sind. Da hierbei auch der Charakter des Ersatzsignals noch durchgängig erkennbar sein soll, dürfen die beim Sprechen häufig auftretenden Pausen nicht zur Unhörbarkeit des Ersatzsignals führen. In einigen Geräten sorgt daher eine automatische Pausenauffüllung (silence bridging) durch eine Erhöhung der Steuerspan-

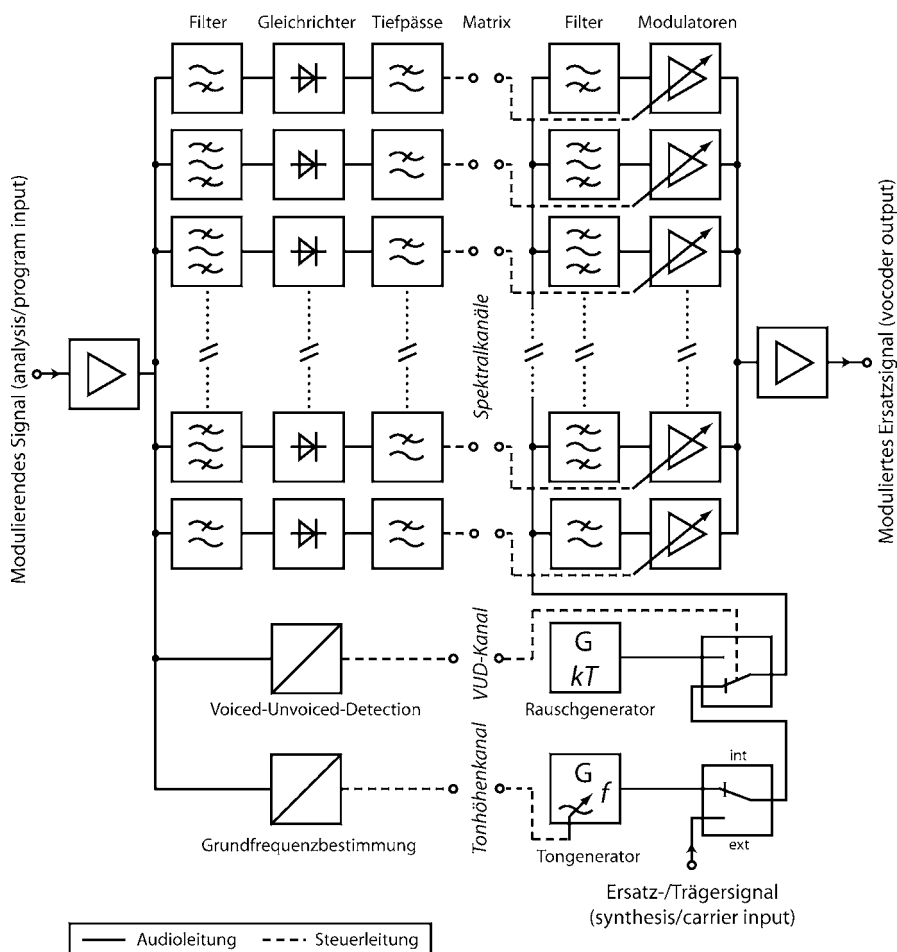


Abb. 13.22 Blockschaltbild des Vocoder

nungen dafür, dass das Ersatzsignal auch bei Sprechpausen den Syntheseteil des Vocoders passieren kann (nicht abgebildet).

Aufgrund der Vielfalt an Vocoder-Bauformen sind sowohl die Bedienungsmöglichkeiten als auch deren Bezeichnungen uneinheitlich. Es können aber folgende Parameter-Gruppen unterschieden werden (mit ggf. abweichenden Bezeichnungen): Eingangspegelsteller je für Sprach- und Ersatzsignal (input level) sowie ggf. Verhältnis von stimmlosem und stimmhaftem Ersatzsignal (input signal balance); Schwellwertregler für VUD (VUD threshold, sibilance level) und Sprache-Pause-Erkennung (speech detector threshold); Regler für Frequenz, Frequenzbereich, Wellenform und ggf. Modulation des VCO (tune, pitch, octave range, waveform, vibrato speed, vibrato depth); Ausgangspegel der einzelnen Spektralkanäle (channel level); Pegel des Pausenfüllsignals, ggf. getrennt für einzelne Spektralkanalgruppen (silence bridging adjust, pausefilling); die Matrix für die Zuordnung der Spektralkanäle (filter bank patch, channel envelope); Ausgangs-Mischfeld für die verarbeiteten bzw. anfallenden Signaltypen: Sprachsignal (microphone/analysis signal level), ggf. analysegefiltertes Sprachsignal spektralkanalweise regelbar (speech addition, multfilter), Ersatzsignal (external/synthesis/replacement signal level), Vocodersignal (vocoder level); ggf. Regler für Filter und nachgeschaltete Effekte (z. B. Chorus).

Die Wirkung der Klangfarbenaufprägung ermöglicht es, v.a. gleichmäßige Klänge wie z. B. Streicherakkorde, Synthesizertöne oder Windgeräusche zum Sprechen (oder bei Tonhöhenänderung zum Singen) zu bringen. Der klassische Vocoder-Effekt ist die Roboterstimme. Durch die zahlreichen Einstellmöglichkeiten, die Wahl der eingespeisten Audioinhalte und die Zuordnung der Steuerspannungen ergeben sich vielfältige Möglichkeiten, Klänge bis zur Unkenntlichkeit zu verfremden. Für einen Überblick über Technik und Einsatzgebiete des Vocoders vgl. Buder (1978a, 1978b). Moderne softwarebasierte Realisationen von Vocodern bieten zusätzlich Visualisierungen, erweiterte Einstellmöglichkeiten und Funktionen (z. B. spektrale Modulation, Morphing, synthetische Klangerzeugung) sowie spektrale Auflösungen von bis zu 1024 Frequenzbändern (vgl. z. B. Haas und Sippel 2004). Eine weiteres Gerät, das eine Aufprägung von sprachlichen Spektralverläufen auf beliebiges Tonmaterial ermöglicht, ist die elektroakustisch arbeitende Talkbox.

13.2.20 Mehrfach-Bearbeitung und komplexe Algorithmen

Moderne Stand-alone-Geräte vereinen häufig verschiedene Audibearbeitungsmittel im Hinblick auf bestimmte Anwendungsfälle oder bestimmte Klangquellen. Typische Beispiele sind im Aufnahmeweg einsetzbare Kanalzüge (channel strips), die in der Regel Vorverstärker, Filter, Equalizer, Regelverstärker und AD-Wandler beinhalten, für Raumeffekte ausgelegte Multieffektgeräte mit Dynamik-, Delay- und Hallprogrammen oder Voice-Prozessoren, die umfangreiche Dynamik-, Tonhöhen- und Klangfarbenbearbeitungen sowie Delay- und Halleffekte ermöglichen. Innovative Verfahren ergeben sich nicht nur durch die Kombination traditioneller Audibearbeitungsmittel, sondern auch durch neue Effekte. So lassen sich in einem sog.

Multiple Resonance Filter Array acht schmalbandige Filter über Hüllkurvengeneratoren von einem Patterngenerator steuern, wodurch eine komplexe klangfarbliche Rhythmisierung möglich wird.

In integrierten Audioproduktionsumgebungen können beliebige Plug-ins kombiniert eingesetzt werden, die sich allerdings die vorhandene Rechenkapazität teilen müssen. Daher sind in diesem Bereich eher klanglich spezialisierte Effekte zu finden. Insbesondere die Anwendung von Frequenzbereichsverfahren ermöglicht die Kreation innovativer Klangbearbeitungsmittel, etwa des Spektral Delay, das frequenzselektive Verzögerungseffekte ermöglicht (vgl. Haas, Clelland & Mandell 2004). Die integrierte Analyse psychoakustisch relevanter Signalmerkmale, wie sie z. B. die Tonhöhenbestimmung für die automatische Tonhöhenkorrektur darstellt, wird künftig weitere Gestaltungsmöglichkeiten eröffnen, etwa die getrennte Bearbeitung von tonhaltigen und geräuschhaften Signalanteilen. Mit einem solchen „intelligenten“ Filter lässt sich z. B. bei einer Flöte das Verhältnis von Anblasgeräusch und Ton regeln. Schließlich gibt es eigenständige Software-Anwendungen, die verschiedene Funktionalitäten (z. B. Synthesizer, Sampler, Effektgerät und Sequenzer) integrieren, sowie Tools für spezielle Anwendungsbereiche (z. B. Sound Design).

13.3 Klangrestauration

Als Restauration bezeichnet man die Entfernung oder Minderung von mit Störungen behaftetem Audiomaterial. Hierbei kann es sich sowohl um historische Aufnahmen als auch um aktuelles Tonmaterial handeln. Bei historischem Material sind die Ursachen für Störungen häufig beschädigte Medien oder Probleme durch die bei der Aufzeichnung verwendeten Aufnahmetechnologien. Die Restauration von alten Aufnahmen ist ein sehr kleines Betätigungsfeld für Spezialisten, aber auch im Studioalltag hat man es oft mit Störungen zu tun, die mit Standardwerkzeugen entfernt werden können.

13.3.1 Bearbeitungswerkzeuge

Zur Bearbeitung kommen entweder externe, DSP-basierte Verfahren zum Einsatz, immer häufiger aber Software-Lösungen für Audio Workstations. Die Algorithmen der meisten Restaurationswerkzeuge bestehen aus einem Funktionsblock, mit dem die Störung detektiert wird, und einem Block, der die detektierte Störung unterdrückt. Es gibt auch Werkzeuge, bei denen die Lokalisation der Störung manuell erfolgen muss. Restaurationswerkzeuge benötigen auf Grund der komplexen Algorithmen sehr viel Prozessorleistung. Häufig verfügen sie über eine Audition-Funktion, über die man den entfernten Störanteil abhören kann. So lässt sich beurteilen, in welchem Umfang auch schon Teile des Nutzsignals entfernt wurden. Es gibt aber auch Störungen, für die bis heute keine Restaurationsalgorithmen entwickelt wur-

den. Entweder sind die Störungen zu komplex oder es gibt noch keine theoretischen Ansatz für eine Restaurierung. Beispiele sind Artefakte durch Mehrfachkodierung oder die Kompensation von Gleichlaufschwankungen bei Bandgeräten.

13.3.2 Klassifizierung von Störungen

Um das richtige Werkzeug für eine Störunterdrückung auszuwählen, muss zunächst der Typ der Störung analysiert werden. Ein Unterscheidungskriterium ist die Frequenz des Störsignals. Es gibt tonale Störungen und Störungen, die sich über ein breites Spektrum verteilen. Ein weiteres Kriterium ist die Stördauer. Hier lassen sich kurze Impulsstörungen von Störungen mit kontinuierlichem Charakter unterscheiden. Sogenannte Bursts oder Cluster sind dabei impulshafte Störungen, die in sehr kurzen Zeitabständen auftreten. Ein weiteres Kriterium ist die Unterscheidung zwischen Störungen, die in der analogen oder in der digitalen Ebene auftreten. Und schließlich kann man unterscheiden zwischen Störungen, die technisch bedingt sind oder Störungen durch andere äußere Einflüsse, wie z. B. Handyklingeln während einer Konzertaufnahme.

13.3.3 Rauschen

Eine der am häufigsten auftretenden Störungen ist das Rauschen. Rauschen ist ein spektral breitbandiges Störsignal mit stochastischem Zeitverlauf (Kap. 1.2.3). Es entsteht bereits im Mikrofon, z. B. durch thermische Bewegung der Luftmoleküle. In der Praxis hat man es jedoch in der Regel mit Rauschen zu tun, das durch elektronische Schaltungen oder analoge Aufzeichnungsmedien verursacht wird.

Bei analogen Systemen, etwa bei analogen Magnetbandgeräten, setzte man dynamische Multiband-Kompressor-Expander-Kombinationen, sog. Kompander zur Rauschunterdrückung ein (Abschn. 13.2.6 und Dickreiter 1990:36ff.). Im Gegensatz dazu arbeiten moderne De-Noiser oder De-Hisser nach der Methode der spektralen Subtraktion. Dabei wird das Spektrum in mehrere Hundert Bänder zerlegt und der Betrag des geschätzten Rauschspektrums vom Gesamtspektrum abgezogen (Boll 1979). Ein prinzipbedingtes Problem der spektralen Subtraktion sind „Musical Tones“, die als Artefakte der eingesetzten Filter kurzzeitige tonale Anteile, eine Art „Klingeln“ erzeugen können. Es existieren verschiedene Lösungsansätze, um diese Artefakte zu unterdrücken (Vary u. Martin 2006). Bei extremen Einstellungen werden diese Musical Tones jedoch hörbar.

In der Praxis ist die Frequenzverteilung des Rauschens nie gleichmäßig. Um das Rauschen aus dem Signal zu entfernen ist daher eine individuelle Ermittlung der spektralen Verteilung des sogenannten Rauschteppichs oder Rauschbodens erforderlich. Dies erfolgt entweder automatisch oder durch einen „Fingerabdruck“, wenn im gestörtem Audiomaterial eine Stelle verfügbar ist, in der ausschließlich das Rau-

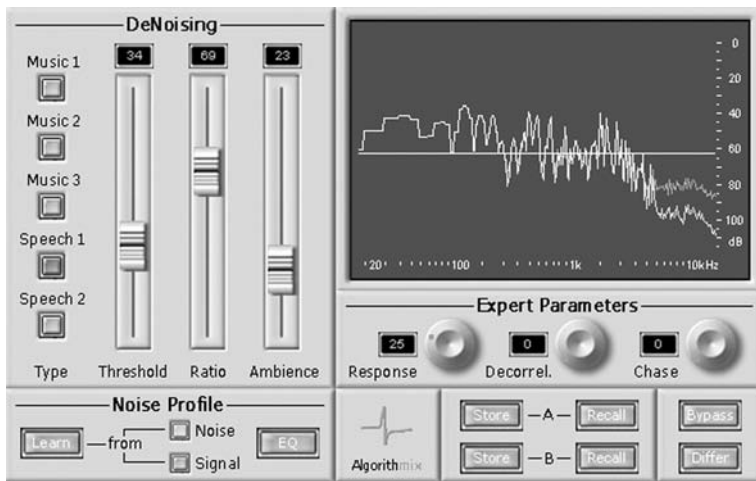


Abb. 13.23 Typischer De-Noiser mit Darstellung der Frequenzverteilung des Nutzsignals und des Rauschbodens

schen vorkommt. Wenn die vom Algorithmus erwartete Länge des zu analysierenden Bereichs nicht ausreicht, setzt man in der Praxis eine Schleife um die Stelle mit dem Rauschstörsignal. Die Analyse eines Fingerabdrucks liefert im Regelfall bessere Ergebnisse als eine automatische Abschätzung des Rauschbodens. Dies gilt insbesondere für Material mit großer Dynamik. Für modulierte Rauschen, wie es bei alten Magnetbandaufzeichnungen vorkommt, sind spezielle Werkzeuge verfügbar und auch notwendig.

Die Ausführung der De-Noiser fällt immer sehr ähnlich aus. Über den Parameter Threshold setzt man die Schwelle für den Rauschpegel, ab der das Restaurationswerkzeug eingreift. Die Stärke der Rauschunterdrückung lässt sich mit dem Parameter Reduction, Depth, manchmal auch als Ratio bezeichnet, einstellen. Professionelle De-Noiser gestatten häufig auch noch die manuelle Korrektur des ermittelten Rauschbodens über einen Equalizer. Durch die Rauschunterdrückung wird in mehr oder weniger großem Umfang immer auch das Nutzsignal beeinträchtigt, insbesondere im Hinblick auf den wahrgenommenen Raumanteil oder die Helligkeit des Tonmaterials. Um den Raumanteil von der Bearbeitung auszuschließen, gibt es meistens einen Ambience-Parameter, mit dem man bestimmen kann, wie groß dieser Anteil sein soll. Auch die Brillanz bei Sprache kann durch den Entrauschprozess gemindert werden. Einige De-Noiser bieten auch hier Parameter, um diese Sprachkomponenten von der Bearbeitung mehr oder weniger auszuschließen.

Ein typisches Problem bei historischem Material ist der scheinbare Verlust an hochfrequenten Anteilen nach der Rauschunterdrückung im Vergleich zum Originalmaterial. Auch wenn das Originalmaterial in den hohen Frequenzbereichen meist kaum oder gar kein Nutzsignal enthält, ist offensichtlich in der Empfindung eine klare Trennung zwischen Nutz- und Störsignal kaum möglich, so dass eine Entfernung des Rauschen im oberen Frequenzbereich als Frequenzbeschneidung wahrgenommen

wird. Da der Einsatz eines Filters lediglich das Rauschen wieder anheben würde, bietet sich der Einsatz eines Werkzeugs zur Generierung von harmonischen Obertönen an (Enhancer, vgl. Abschn. 13.2.9). Die Grenzfrequenz, ab der die Oberwellen erzeugt werden, ist vom Material abhängig und liegt in der Regel oberhalb von 3 kHz.

13.3.4 Scratches, Clicks und Crackles

Bei Impulsstörungen ist eine Klassifikation der Störung besonders wichtig, um das richtige Restaurationswerkzeug auszuwählen. Auch wenn sie sich analoge und digitale Störungen häufig mit den gleichen Werkzeugen bearbeiten lassen, weisen digitale Artefakte bestimmte Besonderheiten auf und sollen daher in Abschn. 13.3.11 separat behandelt werden.

Mechanische Beschädigungen oder Staubeinschlüsse bei Aufzeichnungs- und Wiedergabemedien wie Walzen, Schellack-Platten oder Vinyl-Platten sind in der Regel die Ursache für Impulsstörungen in der analogen Ebene. Bei Schallplatten lassen sich durch das Nassabspielen Probleme durch Staubeinschlüsse in den Rillen wirkungsvoll mindern. Hierbei wird über einen mit Flüssigkeit (Alkohol und destilliertes Wasser) getränkten Schwamm ein Flüssigkeitsfilm auf die Schallplatte aufgetragen. Häufig sind historische Medien aber bereits überspielt, die Originale zu stark beschädigt, oder die Medien sind nicht mehr verfügbar, so dass Restaurationswerkzeuge eingesetzt werden müssen.

Impulsstörungen in der analogen Ebene lassen sich nach Häufigkeit, Dauer und Pegel klassifizieren, die Bezeichnungen für die unterschiedlichen Störungen sind nicht einheitlich. Lange Impulsstörungen von einigen Millisekunden mit hohen Amplituden nennt man Scratches. Mit Clicks bezeichnet man dagegen Störungen mittleren Pegels und mittlerer Dauer. Kurze Störungen mit geringen Pegeln und einer sehr hohen Häufigkeit von mehr als einigen Tausend Ereignissen pro Sekunde bezeichnet man als Crackles. Die Anzahl der Störereignisse ist hier so hoch, dass zum Teil keine Einzelstörungen mehr wahrgenommen werden, sondern die Störung einen tonalen Charakter bekommt. Bei Schallplatten entstehen Scratches und Clicks durch Kratzer auf der Oberfläche. Crackles dagegen rühren im Wesentlichen von den bereits erwähnten Staubeinschlüssen her. Es gibt Restaurationswerkzeuge, die alle drei Störtypen abdecken, aber auch Werkzeuge, die für einen der Störtypen optimiert sind.

Entsprechend des Störtyps werden die Werkzeuge in der Regel als De-Scratcher, De-Clicker oder De-Crackler bezeichnet. Die Restaurationswerkzeuge entfernen die Impulsstörung, die entstehende Lücke wird über eine Interpolation aufgefüllt. Bei Clicks und Scratches kann nach der Bearbeitung eine tieffrequente Störung als Artefakt auftreten. Viele Werkzeuge verfügen daher über eine Funktion, die diesen „Plop“ mindert. Neben dem Einsatzschwellwert (Threshold) und der Höhe der Störungsminderung (Reduction) verfügen die Werkzeuge über weitere Optionen, wie spezielle Betriebsarten für bestimmte Störcharakteristika (z. B. für Schellack-Plat-

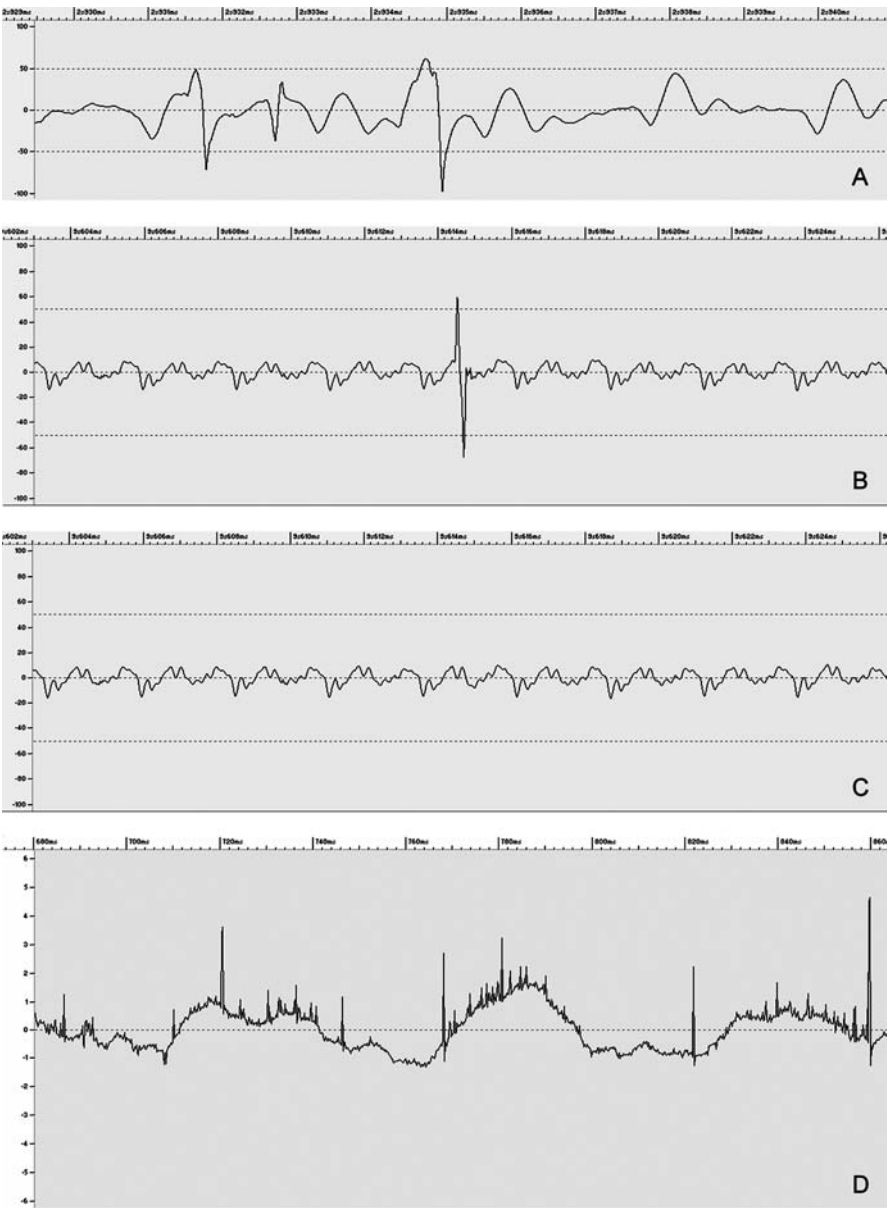


Abb. 13.24 A: Tonsignal von einer Vinylschallplatte mit starken Slices. B: Signal mit einem deutlich wahrnehmbaren Click. C: Signal nach der Bearbeitung durch den De-Clicker. D: Crackles zusammen mit einigen Clicks

ten). Manche De-Scratcher bieten zudem auch die Möglichkeit einen Fingerabdruck der Störung zu nehmen.

Bei Schallplatten treten in der Regel Crackles, Clicks und Scratches gemeinsam auf. Es müssen also entweder mehrere Bearbeitungsdurchgänge erfolgen oder man schaltet die Werkzeuge hintereinander. Als erstes sollten die Störungen mit großen Amplituden und großer Dauer entfernt werden; es empfiehlt sich die Reihenfolge De-Scratcher, De-Clicker und am Ende der Bearbeitungskette den De-Crackler.

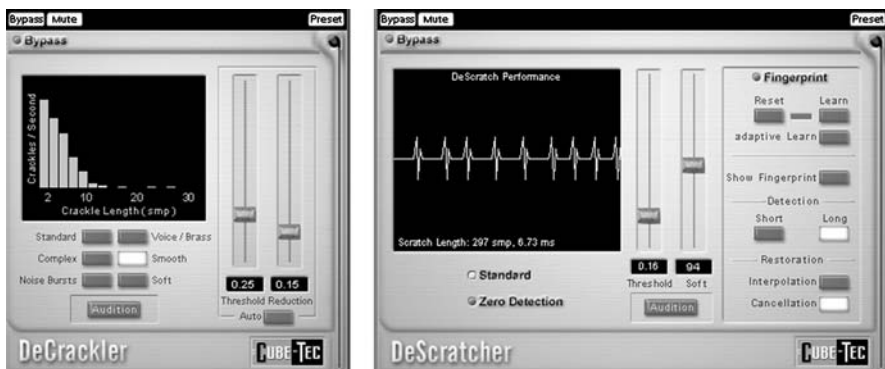


Abb. 13.25 Links: De-Crackler mit einer Anzeige der statistischen Verteilung der Störlängen in Form eines Balkendiagramms. Rechts: De-Scratcher mit grafischer Ausgabe der Zeitfunktion des entfernten Signals

13.3.5 Frequenzgangsbeschnidung

Ein weiteres häufiges Problem bei historischem Tonmaterial sind Frequenzgangsbeschnidungen, die einerseits durch technische Grenzen der Aufzeichnungsmedien bedingt sind, aber auch als Resultat einer bewusst eingesetzten Entzerrung auftreten können. So treten bei historischem Klangmaterial häufig Beschnidungen auf, die nicht allein durch die technischen Bedingungen erklärbar sind. Da eine Kompensation mit Filtern nur das Rauschen anheben würde, wäre auch hier nur der Einsatz von Enhancer-Werkzeugen zur Erzeugung diskreter Obertöne sinnvoll. Im Sinne einer Erhaltung des ursprünglichen Klangs, insbesondere bei historischen Dokumenten, wird allerdings häufig darauf verzichtet, den Klang „aufzufrischen“ und es bleibt bei der Entfernung von Störungen.

13.3.6 Pop-Geräusche

Pop-Geräusche sind tieffrequente Störungen, die meistens durch Impulslaute beim Sprechen oder beim Gesang bei kurzen Mikrofonabständen verursacht werden. Wenn Maßnahmen zum Popschutz (Kap. 7.7.2) nicht eingesetzt wurden, können sie nachträglich durch ein Hochpassfilter unterdrückt werden. Es wird bevorzugt mit dynamischer Kontrolle eingesetzt, um nur die Passagen mit entsprechend hoher Amplitude zu bearbeiten. Da bei dem De-Click-Prozess ähnliche Störungen entstehen können, gibt es auch De-Clicker, die über Hochpassfilter zur Entfernung von Pop-Geräuschen verfügen. Ein weiteres Pop-Geräusch tritt in Form von tieffrequenten Rauschimpulsen bei der Abtastung von Lichttonfilm auf. Auch hier wird ein spezielles, als De-Pop bezeichnetes Werkzeug angeboten.

13.3.7 Übersteuerungen und Verzerrungen

Übersteuerungen (engl.: Clips) können eine vollständige Mischung betreffen (Abb. 13.26), oder aber, insbesondere bei Mehrspuraufnahmen, bereits vor der Mischung in einem der Eingangskanäle aufgetreten sein. Im ersten Fall können sie wirkungsvoll beseitigt werden, indem der Pegel abgesenkt und der durch die Übersteuerung beeinträchtigte Bereich durch eine Interpolation restauriert wird. Ein De-Clipper kann solche Passagen im Signal leicht detektieren und korrigieren. Für eine Übersteuerung von einzelnen Instrumenten in einer Mischung kann ein übersteuertes Mikrofon, ein übersteuertes Aufnahmegerät oder Probleme beim Übertragungsweg

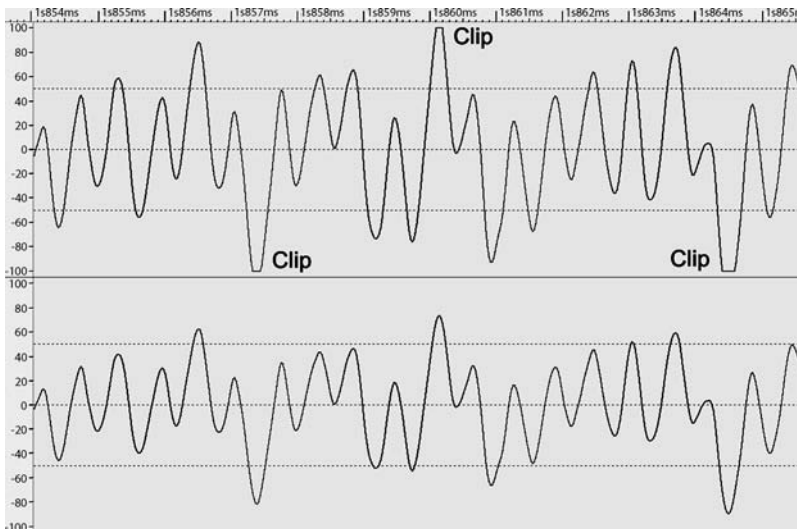


Abb. 13.26 Signal mit Clips auf der oberen Spur. Unten das durch Interpolation restaurierte Signal

(z. B. Drahtlosstrecke) verantwortlich sein. In solchen Fällen bleibt ein De-Clipper weitgehend wirkungslos, eine Minderung der Störung kann allerdings durch Hintereinanderschaltung von zwei oder drei De-Cracklern erzielt werden, die in ihren Schwellwerten und Störreduzierungen gestuft eingestellt werden.

13.3.8 Netz-Störungen

Die Wechselstromversorgung kann zwei unterschiedliche Störtypen hervorrufen. Dies ist zum Einen ein Brummen, z. B. verursacht durch Masseschleifen oder fehlerhafte Schirmung, sowie ein Buzz, eine periodische, obertonreiche Impulsstörung, wie sie durch Dimmer (Phasenanschnittsteuerungen) oder Lichtmischer verursacht wird. Die Grundwelle der Störung, die je nach Wechselspannungsversorgung bei 50 oder 60 Hz liegt, ist häufig schwächer als die Oberwellen.



Abb. 13.27 De-Buzzer zum Entfernen von Brumm- und Dimmer-Störungen

Der naheliegende Lösungsansatz eines sehr steilflankigen Kammfilters ist problematisch, weil die einzelnen Harmonischen unterschiedlich stark vertreten sind und die Frequenz der Wechselspannung, etwa bedingt durch Gleichlaufschwankungen des Aufzeichnungsmediums, geringfügig driften kann. Darüberhinaus überlagern sich Nutz- und Störsignal im Spektrum. Moderne Algorithmen bearbeiten die Harmonischen unabhängig und individuell um so auch eine Beeinträchtigung des Nutzsignals zu vermeiden.

13.3.9 Azimuthfehler

Azimuthfehler bei Bandaufnahmen können durch unterschiedlich justierte Tonköpfe bei der Aufnahme und der Wiedergabe entstehen. Es handelt sich um eine kleine zeitliche Verschiebung, meist nur von wenigen Samples zwischen den Audiokanälen, begleitet von einem Höhenverlust. Ähnliche Laufzeit- bzw. Phasenverschiebungen können auch durch andere Mechanismen in der analogen Ebene entstehen. Durch die Phasenverschiebung wird das Stereobild hörbar beeinträchtigt und es kann bei der Addition der Kanäle zu Kammfiltereffekten kommen.

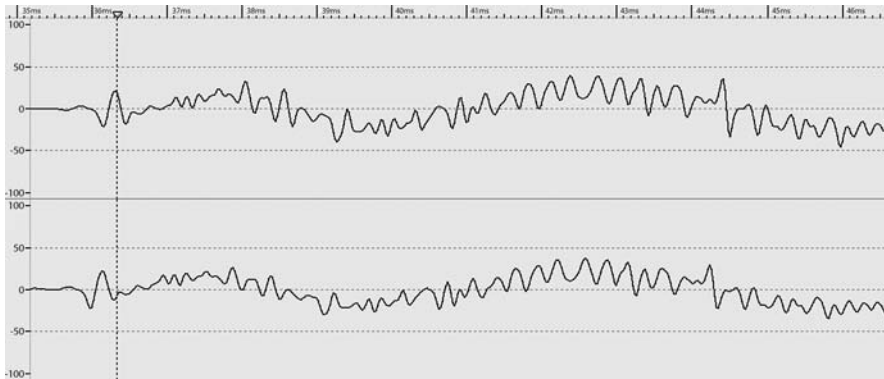


Abb. 13.28 Azimuthfehler einer Bandaufnahme



Abb. 13.29 De-Azimuth Restaurationswerkzeug

Mit einem De-Azimuth-Werkzeug kann die Laufzeit- bzw. Phasenverschiebung erkannt und kompensiert werden. Die Kompensation kann entweder automatisch oder manuell erfolgen. Moderne Restaurationsverfahren sind in der Lage, Phasenver-

schiebungen bis zu einem hundertstel Sample zu erkennen und durch internes Oversampling auszugleichen.

13.3.10 Drop-Outs

Als Drop-Outs werden kurzzeitige Pegel einbrüche bezeichnet, die häufig von einem mehr oder weniger starken Höhenverlust begleitet werden. Ursache kann eine lokale Ablösung der Magnetschicht bei älteren Magnetbändern sein, ein schlechter Band/Kopfkontakt durch lokale Bandverschmutzung, sowie Knicke oder schlechte Klebestellen auf dem Band. Automatische Restaurationswerkzeuge sind für dieses Problem nicht verfügbar, allerdings können Drop-Outs häufig sehr gut manuell bearbeitet werden, indem der Pegel im Drop-Out-Bereich in einem Audioeditor lokal angehoben wird, ggf. auch mit einer lokalen Bearbeitung des hochfrequenten Anteils.

13.3.11 Digitale Störungen

Nicht nur historische Aufnahmen können mit Störungen behaftet sein. Auch in der digitalen Ebene können Störungen auftreten, die in der Regel impulsartig sind oder gebündelt in sog. Clustern auftreten. Einer der häufigsten Ursachen sind Word-clock-, bzw. Synchron-Störungen, die als sporadische Clicks auftreten und sich mit einem Standard-De-Clicker meist gut unterdrücken lassen. Bursts oder Cluster-Störungen treten u. a. bei Problemen mit den digitalen Audio-Schnittstellen auf, und können durch Fehlanpassung, ungeeignete Kabel, Jitter oder falsche Signalamplituden bedingt sein (Kap. 18). Aufgrund ihrer langen Stördauer lassen sie sich meist nur abmildern. Als Werkzeuge kommen je nach Störampplitude Restaurationswerkzeuge für Impulsstörungen wie De-Clicker zum Einsatz, aber auch Werkzeuge, die in der spektralen Ebene arbeiten.

Störungen können auch durch die Editierung des Audiomaterials verursacht werden, so z. B. bei Audioschnitten, die nicht im Nulldurchgang durchgeführt wurden oder bei denen auf ein Crossfade gänzlich verzichtet wurde. Hier treten ebenfalls mehr oder weniger starke Clicks auf, die sich mit einem De-Clicker reduzieren lassen. In professionellen Audioeditoren sind häufig auch manuelle De-Clicker integriert, bei denen ein Bereich um die Störung von Hand selektiert werden muss. Der selektierte Bereich muss hier groß genug gewählt werden, um eine effektive Interpolation zu ermöglichen. Gerade im digitalen Bereich gibt es jedoch eine ganze Reihe von selten auftretenden Störungen wie Block-Wiederholungen, Interpolationen durch Fehlerkorrekturen oder Artefakte durch Audio-Codecs, für die es keine Detektions- und Bearbeitungslösungen gibt.

13.3.12 Störgeräusche

Neben diesen elektronischen Störungen gibt es auch akustische Störungen, die ungewollt in eine Aufnahme geraten, wie Husten, Räuspern, Handyklingeln, Geräusche durch Umblättern der Partitur, Knarren des Bühnenholzbodens und vieles andere. Hierfür bieten einige Hersteller Bearbeitungswerkzeuge an, die in der Lage sind, Störungen durch Bearbeitung im Spektrum zu entfernen. Die Stördetektion erfolgt manuell in einem Spektrogramm. Das Spektrogramm (in der Phonetik auch als Sonagramm bezeichnet), ist eine gleitende Kurzzeit-FFT, die mit Blockgrößen in der Größenordnung von 1024 Samples über die Dauer des zu analysierenden Audiosignals durchgeführt wird. Das Ergebnis wird meist mit vertikaler Frequenzachse und horizontaler Zeitachse dargestellt, die Amplitude der jeweiligen Frequenz wird durch verschiedene Farben oder verschiedene Helligkeitswerte dargestellt. Durch Veränderung der Blockgröße lässt sich entweder die Frequenzauflösung auf Kosten der Zeitauflösung oder die Zeitauflösung auf Kosten der Frequenzauflösung verändern, wodurch Störkomponenten optisch deutlicher sichtbar gemacht werden können. In einem solchen Diagramm lässt sich mit einiger Übung eine Störung leicht erkennen. Tonale Störungen zeichnen sich als horizontale Linien ab, Breitbandstörungen nehmen einen größeren vertikalen Bereich ein.

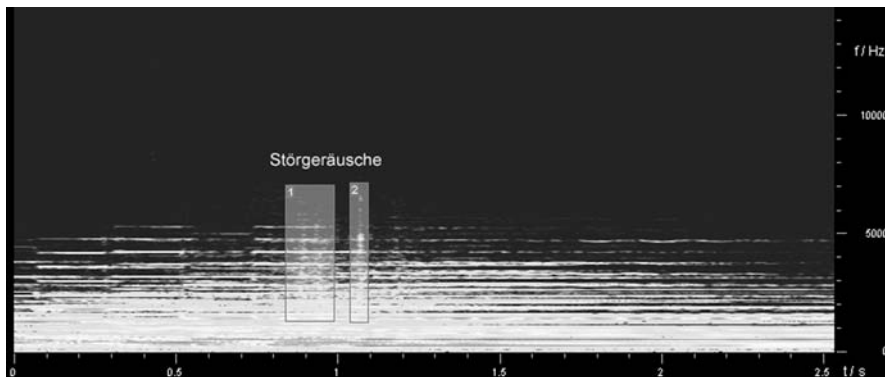


Abb. 13.30 Spektrogramm mit Rascheln, verursacht vom Umblättern einer Partitur

Die Störbereiche im Spektrum werden mit der Maus markiert, mit Hilfe des Spektrums um die entfernten Bereiche herum erfolgt eine Interpolation des Spektrums im Störbereich. Häufig sind mehrere Bearbeitungsvorgänge nötig. Als Bearbeitungsartefakt können „Löcher“ (ähnlich wie Drop-Outs) oder tonale Komponente hörbar werden. In diesem Fall hilft eine geringere Störunterdrückung oder die Bearbeitung eines kleineren spektralen Bereichs.

Für bestimmte technische Geräusche gibt es spezielle Restaurationswerkzeuge, z. B. für das Unterdrücken von Kamera-Transport- und Zoom-Geräuschen. Sie arbeiten ähnlich wie ein De-Noiser, allerdings mit einer an den veränderlichen Störpegel angepassten, dynamischen Unterdrückung.

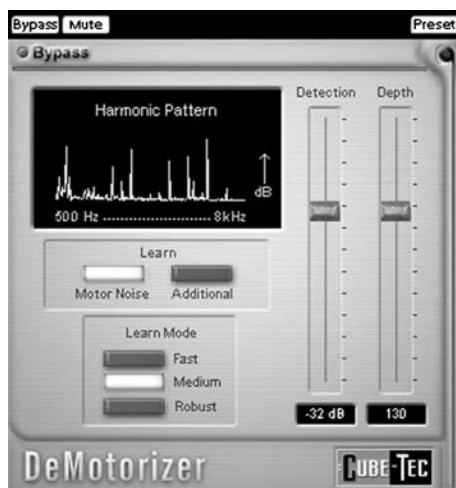


Abb. 13.31 De-Motorizer – ein spezielles Werkzeug zum Entfernen von typischen Film- und Video-Kamerageräuschen

13.3.13 Praxis

Die Restauration von Audiomaterial ist immer eine Suche nach dem bestem Kompromiss, denn die Bearbeitung ist fast immer von Artefakten, also von ungewollten Nebenerscheinungen begleitet. Diese Artefakte können sich sowohl als neue Störungen bemerkbar machen oder als Beeinträchtigung des Originalmaterials, etwa durch Frequenzgangsbeschränkung. Aus diesem Grund müssen Restaurationswerkzeuge so eingesetzt werden dass Störungen ausreichend unterdrückt werden und Artefakte dabei möglichst unhörbar bleiben.

Ein wichtiges Anliegen bei historischem Material kann es sein, den Klangcharakter der Zeit, in der das Audiomaterial aufgezeichnet wurde, und damit den Eindruck von Authentizität nicht zu beeinträchtigen. Bei Audiomaterial aus dem letzten Drittel des 20. Jahrhundert überwiegt dagegen der Wunsch nach einem Klang nach heutigen Maßstäben. Diese Adaption des Klangcharakters bezeichnet man im allgemeinen als Re-Mastering.

Die Algorithmen in Restaurations-Werkzeugen sind von Hersteller zu Hersteller unterschiedlich, auch wenn Funktionalität und Bedienung auf den ersten Blick häufig ähnlich erscheinen. Somit können verschiedene Algorithmen bei der gleichen Störung sehr unterschiedliche Wirkungen erzielen. Restaurations-Experten verfügen daher über eine große Auswahl an Werkzeugen, die alternativ ausprobiert werden können. Gelegentlich lassen sich auch durch eine Hintereinanderschaltung von Werkzeugen die Resultate verbessern.

13.4 Funktionen und ästhetische Ziele

Werden akustische Darbietungen, insbesondere Musik, technisch bzw. medial übertragen, werden sie in mehrfacher Weise transformiert:

- akustisch (Original- und reproduziertes Schallfeld sind nicht identisch)
- klanglich (das auditive Perzept ist ein anderes)
- interpretatorisch (in die elektroakustische Übertragung fließen künstlerische, ästhetische und technische Entscheidungen Dritter ein)
- räumlich (die Wiedergabe erfolgt in einem anderen Raum)
- zeitlich (meist werden Aufzeichnungen gehört)
- kontextuell (akustische, optische, motivationale, emotionale, kognitive und soziale Bedingungen der Rezeption sind verändert)

Dass die mediale Transformation das Erleben von Musik beeinflusst, zeigte schon Reinecke (1978). Die Audiotbearbeitung ist dabei insoweit eine wesentliche Einflussgröße, als sie sich – im Gegensatz zu anderen Transformationsfaktoren – auf den technisch manifesten Audioinhalt auswirkt und daher für seine weitere Übertragung weitgehend Bestand hat, während hingegen Übertragungs- und Rezeptionsbedingungen variieren können.

Die Wirkung bestimmter klanggestalterischer Maßnahmen oder Prozesse ist erst seit einigen Jahren Gegenstand wissenschaftlicher Untersuchungen. Dass sich Produktions- und Audiotbearbeitungsmaßnahmen, die sprachinhaltliche, musikalisch-strukturelle oder aufführungsinterpretatorische Eigenschaften verändern (z. B. Aufnahmeleitung oder Tonmontage) auf die Wahrnehmung ebendieser Eigenschaften auswirken, mag selbstverständlich sein. Irrtümer der Musikkritik – eine aufgezeichnete Aufführung wird nach einem Remastering für eine andere gehalten (vgl. Stolla 2004: 11) – und experimentelle Untersuchungen zeigen jedoch, dass die Einschätzung substanzieller Eigenschaften (wenigstens bei Musikinhalten) allein durch klangliche Bearbeitungsmaßnahmen beeinflusst werden kann. So beurteilten Musikprofessoren und -studenten in einem Experiment von Boss (1995) zwei verschiedenen mikrofonierte und nachbearbeitete Aufnahmen derselben kammermusikalischen Darbietung hinsichtlich interpretatorischer Merkmale wie Tempo, Artikulation, Agogik oder Phrasierung deutlich unterschiedlich, was sich auch in der Präferenz der Versionen niederschlug. Verschiedene Abmischungen und klangliche Nachbearbeitungen eines unbekannten Popmusik-Titels beeinflussten ebenfalls nicht nur klangbeschreibende, sondern auch ästhetische, emotionale und strukturbeschreibende Beurteilungsmerkmale sowie die Kaufpräferenz in musikwirtschaftlich relevanter Größenordnung (Maempel 2001). Unter wirkungsästhetischen Gesichtspunkten stellt die Klanggestaltung durch Audiotbearbeitungsmittel also einen bedeutsamen Aspekt der Medienproduktion dar, die auch auf die Beurteilung der Künstler zurückwirken kann. Wiedergabeseitige klanggestalterische Eingriffe durch den Rezipienten zeigten ebenfalls Wirkungen, die weit über die Klangwahrnehmung hinausgehen. So beeinflussten im Auto-Fahrsimulator Lautstärke und Equalisierung des gehörten Musikprogramms Reaktionszeit und Geschwindigkeitswahl des Fahrers (Haack 1990).

13.4.1 Physik und Psychologie

Für ein eingehendes Verständnis der Möglichkeiten und Grenzen der Audiotbearbeitung ist phänomenologisch grundsätzlich zwischen physikalischer und psychologischer Ebene zu unterscheiden. Durch die Audiotbearbeitungsvorgänge werden zunächst nur physikalische Maße des Audiosignals zuverlässig verändert: Amplituden- und Zeitwerte oder andere hieraus konstruierte Signalmaße wie Frequenz und Phasenlage, und zwar zumeist in Abhängigkeit voneinander. Der Klang und die strukturelle Erkennung eines Audiosignals hingegen sind als reiner Wahrnehmungsinhalt (Perzept) ausschließlich psychologisch repräsentiert, und zwar in den grundlegenden (jedoch nicht unabhängigen) Merkmalen Lautstärke, Klangfarbe, Tonhöhe, räumliche Position, Raumeindruck und zeitliche Lage (Abb. 13.32). In mittleren und höheren psychischen Verarbeitungsstufen ergeben sich zahlreiche weitere Inhalte, Merkmale oder Vorgänge: von Aufmerksamkeit über Gestalt- und Objekterkennung bis hin zu Bedeutung, Bewertung, Erinnerung, Gefühl und Handlungsmotivation. Obwohl physikalische und psychologische Merkmale also unterschiedliche Phänomene beschreiben und strukturell inkongruent sind, ist die Vorstellung einer weitgehend unvermittelten Abbildbarkeit der einen Merkmalsstruktur auf die andere weit verbreitet. Schon 1978 versuchte Nitsche, physikalische Maße von Audiosignalen zu konstruieren, die zuverlässig mit Klangfarbeneindrücken korrelieren, was jedoch kaum gelang. Furmann, Hojan, Niewiarowicz & Perz (1990) fanden, dass nur eines von fünf erhobenen perzeptiven Merkmalen (Schärfe) signifikant mit phy-

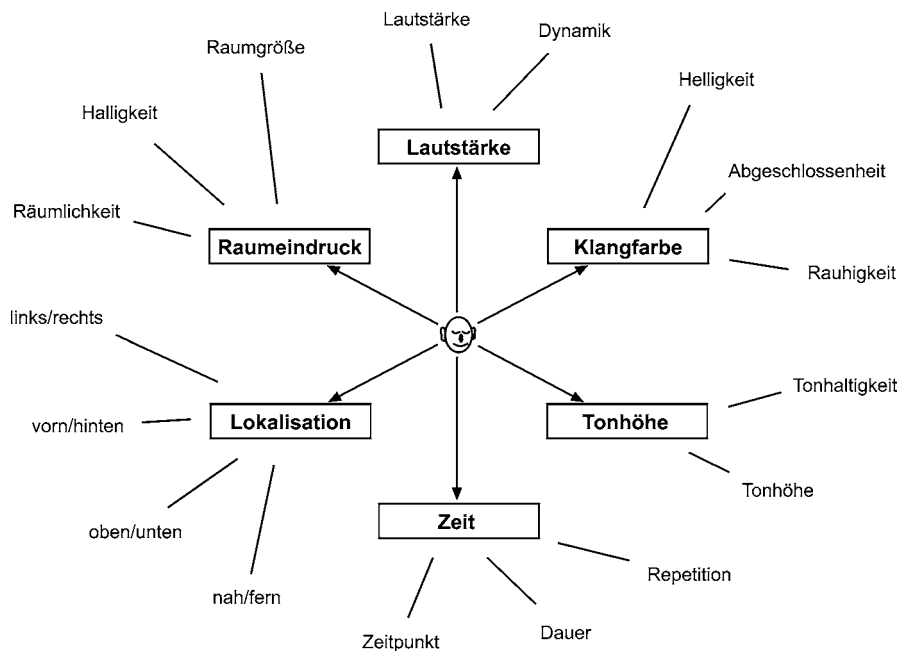


Abb. 13.32 Grundlegende Merkmale der Klangwahrnehmung

sikalischen Merkmalen getesteter Lautsprecher korrelierte. Eine bessere Vorhersagbarkeit einfacher perzeptiver Beurteilungen durch z. T. ex post konstruierte physikalische Maße fanden Lehmann & Wilkens (1980) in einer Untersuchung zur Akustik von Konzertsälen, jedoch auch eine starke Beeinflussung durch die verwendeten Musikstücke und den persönlichen Geschmack.

Im wesentlichen erschweren bzw. verhindern zwei Gründe die Integration der physikalischen und der psychologischen Ebene: Zum einen sind die Beziehungen zwischen den jeweiligen Merkmalen komplex. So kann die Veränderung eines technischen Parameters (z. B. Frequenzspektrum) die Veränderung der Ausprägung mehrerer Klangmerkmale zur Folge haben (z. B. Klangfarbe, Lautheit, Entfernungseindruck), oder zur Veränderung der Ausprägung eines Klangmerkmals (z. B. Entfernungseindruck) müssen mehrere technische Parameter verändert werden (Pegel, Frequenzspektrum, Verhältnis von Direkt- und Diffusschall). Zum anderen ist der Mensch das einzige gültige Messinstrument zur Bestimmung von Wahrnehmungsinhalten, er liefert aber inter- wie intraindividuell nicht so zuverlässige Daten wie eine physikalische Messung, was durch individuelle Konstruktionsvorgänge als Teil des Wahrnehmungsprozesses bedingt ist. Physikalische und perzeptive Messverfahren unterscheiden sich vor allem hinsichtlich der Erfüllung der klassischen Testgütekriterien Validität (Gültigkeit im Sinne von Eignung der verwendeten Variablen für die interessierenden Merkmale sowie von eindeutiger Interpretierbarkeit und Generalisierbarkeit der Messung) und Reliabilität (Zuverlässigkeit im Sinne von Genauigkeit und Reproduzierbarkeit der Messergebnisse). Die Erzielung einer hohen Objektivität (Nachprüfbarkeit und Vermeidung von Versuchsleiter-Subjektivität) hingegen ist, sofern mit standardisierten Instrumenten und transparent gearbeitet wird, bei beiden Ansätzen weitgehend unproblematisch, weswegen der Begriff *subjektiv* für die Beschreibung von Mehr-Probanden-Hörversuchen irreführend ist und deren wissenschaftliche Qualität unnötig negativ attribuiert. Aufgrund der genannten Umstände können prinzipiell keine deterministischen, sondern ausschließlich probabilistische Zusammenhänge zwischen physikalischer und psychologischer Ebene ermittelt werden. Weitgehend eindeutige Beziehungen können nur für wenige grundlegende Merkmale hergestellt werden (Zwicker u. Feldtkeller 1967:1, Fastl 1997). Zudem wurden diese zumeist unter Laborbedingungen (z. B. unbeweglicher Kopf) und für künstliche Audio-Inhalte (z. B. Sinustöne) empirisch ermittelt. Mit der Herstellung solcher – gemessen an der Leistungsfähigkeit des Hörsinns und der auditiven Verarbeitung einfachen – Zusammenhänge befasst sich die klassische Psychoakustik. Produkte höherer psychischer Verarbeitungsstufen wie komplexe Klangmerkmale (z. B. Transparenz), ästhetische Eindrücke oder Sinnerkennung hingegen können bislang überhaupt nicht allgemeingültig auf spezifische physikalische Eigenschaften des akustischen Reizes zurückgeführt werden. Gerade die Beeinflussung dieser höheren menschlichen Wahrnehmungsinhalte ist aber überwiegend das Ziel der Audiotbearbeitung. Der Audio-Produzent kann sich hierbei also kaum auf wissenschaftlich belegte Zusammenhänge stützen, sondern nur auf individuelle Konzepte, Versuch und Irrtum, die eigene Wahrnehmung, Erfahrung und künstlerisch-ästhetische Entscheidung. Doch auch hierfür lassen sich einige allgemeine Wirkungsprinzipien, Ziele und Funktionen formulieren, denn für

grundlegende Wahrnehmungsvorgänge wie Empfindungs-, Aufmerksamkeits- und Erkennungsprozesse können typische, interindividuell ähnliche Funktionsweisen vorausgesetzt werden. Hierzu zählen

- psychoakustische Mechanismen, insbesondere Maskierung und Lokalisation,
- Gesetzmäßigkeiten, z.B. Gesetze der Nähe, der Ähnlichkeit, der Einfachheit, (vgl. Goldstein 2002:190–204) sowie
- Zusammenhänge im Sinne der neuen experimentellen Ästhetik (Berlyne 1971, 1974) zwischen reizbezogenen kollativen Variablen (z. B. Neuartigkeit, Komplexität) und Reaktionsdimensionen (z. B. Aufmerksamkeit, Bewertung).

Viele in der Produktionspraxis berücksichtigte funktionale und ästhetische Gesichtspunkte insbesondere der Klanggestaltung lassen sich auf die genannten Zusammenhänge zurückführen. Sie treten allerdings zumeist mit inhaltlichen Bedeutungen in Wechselwirkung. Dies gilt insbesondere für Musik, sofern man die Erkennung musikalischer Struktur als eine Voraussetzung für das Entstehen musikalischer Bedeutung ansieht. Der Wahrnehmung musikalischer Struktureinheiten liegen Prinzipien der Gestalterkennung zugrunde (Motte-Haber 2005), und der Bottom-up-Prozess (reizgeleitete Informationsverarbeitung) als Teil der Gestalterkennung hängt von klanglichen Eigenschaften ab: Grundsätzlich erhöhen ähnliche physikalische und perzeptive Eigenschaften (cues) von Schall- bzw. Hörereignissen die Wahrscheinlichkeit ihrer Fusion zu einem Auditory Stream (Bregman 1990). Zur Konturierung musikalischer Gestalten tragen auch Spektraldifferenzen bei (Terhardt 1986). So fördert z.B. Klangfarbenhomogenität die Wahrnehmbarkeit einer geschlossenen Melodielinie (Wessel 1979). Nachdem Klang der musikalischen Struktur insoweit inhärent ist (Maempel 2001:21–24), beeinflusst eine selektive Klangbearbeitung immer auch die Wahrnehmung musikalischer Bedeutung.

13.4.2 Gestalten und Interpretation

Eine sinnvolle, d.h. an inhaltlicher Bedeutung orientierte Audiotbearbeitung sorgt für die Erkennbarkeit, Schärfung, Trennung, Zusammenfassung und Gewichtung musikalischer Gestalten oder anderer inhaltlicher Einheiten. Hierzu kann auf alle auditiven Merkmale (vgl. Abb. 13.32) Einfluss genommen werden. Sowohl die Ermittlung der musikalischen bzw. inhaltlichen Intention als auch die konkrete Festlegung von Klangeigenschaften – dasselbe Ziel kann mit verschiedenen Mitteln erreicht werden – ist jeweils ein klarer interpretatorischer Vorgang (neben der Aufführungs-Interpretation die so genannte zweite Interpretation). Audiotbearbeitung ist im Rahmen der Musikübertragung daher nicht nur ein technischer, sondern auch und gerade ein musikalischer Prozess (Schlemm 1997, Maempel 2001). Gleichwohl lassen sich nach Dickreiter (1997:336) räumliche Ausgewogenheit („Symmetrie“) des Klangbilds und Klarheit als übergeordnet gültige klangästhetische Prinzipien ausmachen.

13.4.3 Klangästhetische Maximen

Der Musikübertragung können grundsätzlich drei divergierende Zielsetzungen zugrunde liegen (Terminologie nach Stolla 2004):

- die Reproduktion des physikalischen Schallfelds des Aufführungsorts durch möglichst unverfälschte Signalübertragung (positivistisches Klangideal),
- die Erzeugung eines der Aufführung ähnlichen Musikerlebnisses durch klangliche Beeinflussungen, die eine entsprechende Illusion fördern (illusionistisches Klangideal),
- die interpretatorisch freie Realisation der Partitur ohne zwingenden Aufführungsbezug (medial-autonomes Klangideal).

Diese fundamentale Einteilung ist auf die Übertragung aller auditiven Inhalte anwendbar.

13.4.3.1 Kunstmusik

Nach Stolla (2004) folgen Übertragungen von konzertanter Kunstmusik seit den 1970er Jahren einer illusionistischen Ästhetik: Aufnahmen orientieren sich klanglich am Aufführungsoriginal und zielen auf Natürlichkeit ab, die es jedoch wegen des komplexen Transformationsprozesses medial nicht geben kann (vgl. hierzu auch Bickel 1992 und Schlemm 1997), zumal sie durch die gleichzeitige Forderung nach spieltechnischer Perfektion untergraben wird. Audiotbearbeitung wird im Rahmen des illusionistischen Konzepts mit Blick auf die Ebene des psychischen Erlebens nicht gemäß dem puristischen Ansatz minimiert, sondern gezielt und ggf. deutlich zur Überhöhung klanglicher Eigenschaften eingesetzt. Durch die so verbesserte Plastizität auditiver musikalischer Gestalten und Bedeutungen sollen transformationsbedingt fehlende (z. B. visuelle) Merkmale der Aufführung kompensiert werden.

13.4.3.2 Populäre Musik

Populärmusikalische Genres, Radiokunst und in der Regel Hörspiel folgen hingegen einem medial-autonomen Klangideal. Ein kreativer und innovativer Einsatz von Audiotbearbeitungsmitteln ist hier nicht nur zulässig, sondern geradezu erforderlich. So nimmt populäre Musik zumeist nicht auf vorbestehende oder klanglich klar definierbare Aufführungssituationen Bezug. Sie entsteht nicht vor, sondern in dem technischen Medium, das hier neben der vermittelnden auch eine Produktionsfunktion besitzt. Ungeachtet der Frage, was unter diesen Umständen als Original bezeichnet werden kann, erlaubt die Freiheit, nicht einem Ideal aufführungsbezogener Natürlichkeit genügen zu müssen, eine besonders wirksame Unterstützung der Wahrnehmung musikalischer Strukturen. So wird in der Regel mit starken klanglichen Kontrasten gearbeitet, um den (häufig strukturell stereotypen) musikalischen Gestalten hohe Prägnanz zu verleihen (Maempel 2001).

13.4.4 Aktivierung

Da Audioinhalte häufig werbende oder selbstwerbende Funktionen erfüllen sollen (z. B. Werbespots, Kennmelodien oder Popularmusik), wird in der Regel versucht, ein Mindestmaß aktivierender Eigenschaften sicherzustellen mit dem Ziel, beim Hörer Aufmerksamkeit und Hinwendung zu erzeugen. Die Aktivierung des Hörers kann auf klanglicher Ebene durch eine Erhöhung der Reizenergie (z. B. durch Helligkeit und Lautheit) oder eine Erhöhung der Reizkomplexität (z. B. durch den Einsatz von Effekten und unbekannten Sounds sowie durch schnelle Wechsel von Klangeigenschaften) gefördert werden. In Verbindung mit einem ästhetisch hochkomplexen Bewegtbild (z. B. Videoclip) werden in der Tendenz vornehmlich energiehaltige, ohne Bild hingegen eher komplexe klangliche Gestaltungen bevorzugt (Maempel 2001). Grundsätzlich trägt auch ein geeignetes Verhältnis von Neuem und Vertrautem bei der Soundauswahl und -bearbeitung zu einer positiven Hinwendung des Hörers bei. Das Prinzip ist unter dem Kürzel MAYA (most advanced yet acceptable) bekannt. Ob das aus der Werbepsychologie entlehnte Aktivierungskonzept im heutigen Umfeld der Reizüberflutung uneingeschränkt aufgeht, ist allerdings fraglich: Insbesondere die energetischen Merkmale scheinen weitgehend ausgereizt zu sein. Sie sind zudem in den meisten Fällen durch den Rezipienten beeinflussbar (z. B. Lautstärkewahl), also nicht nachhaltig wirksam. Mit dem Ziel, den Bewegungsdrang des Hörers anzusprechen und ihn damit umfassender zu aktivieren, wird auch auf eine feine mikrorhythmische Gestaltung Wert gelegt. Durch fein abgestimmte Pegel- und spektrale Verhältnisse und den Einsatz von Delayeffekten wenigstens innerhalb der Rhythmus-Sektion kann ein schon strukturell angelegter Groove eines Musikstücks befördert werden.

13.4.5 Assoziation

Audiobearbeitungsmaßnahmen können wie Audioinhalte auch klare eigene Bedeutungen tragen, also selbst Inhalt sein, oder wenigstens ein bestimmtes Assoziationspotenzial besitzen. So lassen sich durch Filterung und Verzerrung Telefon- und Megafonstimmen nachahmen, der Charakter einer Großveranstaltung, eine Clubatmosphäre oder eine sakrale Stimmung können durch den Einsatz entsprechender Halleffekte hergestellt oder verstärkt werden, und die Verwendung oder Simulation historischer Audiotechnik (z. B. Unterlegen von Schellack- oder Vinylplatten-Geräuschen, Filterung) verweist auf eine vergangene Epoche, einen bestimmten Zeitgeist oder ein bestimmtes Lebensgefühl.

13.4.6 Robustheit und Lautheit

Für viele massenmediale Inhalte gewinnt ein weiteres Kriterium zunehmend an Bedeutung. Das Klangbild sollte möglichen Unzulänglichkeiten der medialen Übertragungskette in dem Sinne Rechnung tragen, dass das inhaltlich oder musikalisch Wesentliche deutlich erkennbar übertragen wird. Vor dem Hintergrund der zunehmenden Diversifikation und Portabilität von Wiedergabegeräten (z. B. Autoradio, Laptop, Handy, MP3-Player), muss heute nicht nur von einer Verletzung der optimalen Abhörbedingungen (z. B. Standard-Stereoaufstellung, vgl. Abb. 13.4), sondern auch grundsätzlich von zunehmend ungünstigen Abhörbedingungen (z. B. schmalbandiger Frequenzgang, hoher Klirrfaktor, laute Umgebung) und Hörgewohnheiten (z. B. Nebenbeihören) ausgegangen werden. Diesen Umständen wird oft durch eine stärkere Monokompatibilität, vereinfachte Lokalisationsverteilung, geringere Dynamik bzw. höhere Lautheit, klangfarbliche Aggressivität und Vermeidung der Nutzung der Außenoktaven für musikalisch wesentliche Information begegnet. Sowohl diese *Robustheit des Klangbilds* als auch der Lautheitswettbewerb sind Beispiele für die Einengung potenzieller klangästhetischer Vielfalt auf medial funktionale Varianten (Maempel 2007).

Die klangästhetischen Überlegungen legen oft gegensinnige Bearbeitungsmaßnahmen nahe, so dass hier Prioritäten gesetzt und Kompromisse eingegangen werden müssen. Dies ist auch bei der klangästhetischen Beurteilung zu berücksichtigen, die daher nicht qualitativ punktuell erfolgen, sondern stets die Gesamtheit der potenziell relevanten Aspekte im Blick haben sollte.

Literatur

- Antares Audio Technologies (Hrsg) (2006) Auto-Tune 5 Pitch Correcting Plug-in Owner's Manual. Scotts Valley / CA http://www.pair.com/anttech/downloads/AT5_manual.pdf. Zugriff 16.2.2007
- Axon PE, Gilford CLS, Shorter DEL (1957) Artificial Reverberation. *J Audio Eng Soc* 5/4:218–237
- Baeder KO, Blesser BA (1975) Klangumformung durch Computer. 10. Tonmeistertagung, Köln, S 24–29
- Berlyne DE (1971) *Aesthetics and Psychobiology*. Appleton-Century-Crofts, New York
- Berlyne DE (Hg) (1974). *Studies in the New Experimental Aesthetics. Steps toward an Objective Psychology of Aesthetic Appreciation*. Hemisphere, Washington
- Bernsee SM (2005) Time Stretching And Pitch Shifting of Audio Signals – An Overview. <http://www.dsdimension.com/index.html?timepitch.html>. Zugriff 27.3.2007.
- Bertsch M (1994) Der Trompetendämpfer. Einfluß des Dämpfers auf das akustische Verhalten und die Klangfarbe der Trompete. In: Bildungswerk des Verbands Deutscher Tonmeister (Hrsg) Bericht zur 18. Tonmeistertagung in Karlsruhe, Saur, München, S 95–100.
- Blauert, Jens (1974) *Räumliches Hören*. (Monographien der Nachrichtentechnik). Hirzel, Stuttgart
- Boer K de (1940) Plastische Klangwiedergabe. *Philips' technische Rundschau* 5:107–115.
- Boll S (1979) Suppression of Acoustic Noise in Speech using Spectral Substraction. *IEEE Trans on ASSP* 27:113-120
- Bregman AS (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambridge/MA

- Brunner S, Maempel H-J, Weinzierl S (2007) On the Audibility of Comb Filter Distortions. 122nd AES Convention, Vienna, Austria, Preprint 7047
- Buder D (1978a) Vocoder für Sprachverfremdung und Klangeffekte, Teil 1. *Funkschau* 7:293-297
- Buder D (1978b) Vocoder für Sprachverfremdung und Klangeffekte, Teil 2. *Funkschau* 8:337-341
- Celibidache S (1985) Musik verschwindet: Interview der Autoren mit Sergiu Celibidache 1985. In: Fischer M, Holland D, Rzehulka B (1986) (Hrsg) *Gehörgänge: Zur Ästhetik der musikalischen Aufführung und ihrer technischen Reproduktion*. Peter Kirchheim Verlag, München, S 115–130
- Chowning JM (1971) The Simulation of Moving Sound Sources. *J Audio Eng Soc* 19/1:2-6
- Craven PG (2003) Continuous surround panning for 5-speaker reproduction. 24th AES Conference Banff, Canada
- Culshaw J (1981) Putting the Record Straight. **Secker & Warburg, London**
- Dickreiter M (1997) *Handbuch der Tonstudientechnik*. Bd.1, 6. Aufl, Saur, München
- EMT (1983) Ein Vierteljahrhundert Hall in der Hand des Tonmeisters. *EMT Kurier*, Mai 1983
- Eventide Inc. (Hrsg) (2006) H8000FW Operating manual (for software version 5.0). Little Ferry / NJ http://www.eventide.com/pdfs/h8000fw_um.pdf. Zugriff 16.2.2007
- Fastl H (1997) „The Psychoacoustics of Sound-Quality Evaluation“. *Acustica* 83/5:754-764
- Flanagan JL (1972) *Speech analysis, synthesis, and perception*. 2. Aufl, Springer, Berlin
- Furmann A, Hojan E, Niewiarowicz M, Perz P (1990) On the Correlation between the Subjective Evaluation of Sound and the Objective Evaluation of Acoustic Parameters for a Selected Source. *J Audio Eng Soc* 38/11:837-844.
- Gardner WG (1995) Efficient convolution without input-output delay. *J Audio Eng Soc* 43/3
- Gernemann-Paulsen A, Neubarth K, Schmidt L, Seifert U (2007). Zu den Stufen im Assoziationsmodell. Bericht der 24. Tonmeistertagung, Leipzig, S 378–407
- Gerzon MA (1992). Panpot laws for multispeaker stereo. 92nd AES Convention, Wien, Preprint 3309
- Goldstein EB (2002) *Wahrnehmung*. Heidelberg, 2. Aufl, Spektrum, Berlin
- Gooch, BR (1999) Building on the Magnetophon. In: Daniel ED, Mee CD, Clark MH (Hrsg) *Magnetic Recording. The First 100 Years*. IEEE press, New York, S 87f
- Gould G (1966) Die Zukunftsaussichten der Tonaufzeichnung. In: Fischer M, Holland D, Rzehulka B (1986) (Hrsg) *Gehörgänge: Zur Ästhetik der musikalischen Aufführung und ihrer technischen Reproduktion*. Peter Kirchheim Verlag, München
- Haack S (1990) Wirkungen übertragungstechnischer Faktoren. In: Motte-Haber H, Günther Rötter (Hrsg) *Musikhören beim Autofahren: 8 Forschungsberichte*. Lang, Frankfurt am Main, S 94–125
- Haas J, Sippel S (2004) *VOKATOR Benutzerhandbuch*. Native Instruments Software Synthesis GmbH
- Haas J, Clelland K, Mandell J (2004) *NI-Spektral Delay Benutzerhandbuch*. Native Instruments Software Synthesis GmbH
- Hartmann WM (1978) Flanging and Phasers. *J Audio Eng Soc* 26/6:439-443.
- Hartmann WM (2005) *Signals, Sound, and Sensation*. Corr 5th printing (**Modern acoustics and Signal Processing**). Springer, New York
- Hartwich J, Spikowski G, Theile G (1995) Ein Dynamic-Kompressor mit gehörangepaßten statischen und dynamischen Kennlinien. Bericht der 18. Tonmeistertagung, Karlsruhe, S 72–84.
- Hoeg W, Steinke G (1972) *Stereofonie-Grundlagen*. Verl Technik, Berlin
- Hoenig UG, Neubäcker P (2006) melodyne 3 studio & cre8 handbuch version 3.0. http://www-celemony.com/cms/uploads/media/Manual.MelodyneCre8Studio.3.0.German_03.pdf. Zugriff 16.2.2007
- Jakubowski H (1985) Dynamik – Versuch einer Begriffsbestimmung. In: Bericht der 13. Tonmeistertagung, München, S 296–304
- Jot JM, Chaigne A (1991) Digital delay networks for designing artificial reverberators. 90th AES Convention, Paris, Preprint 3030
- Kuhl W (1958) Über die akustischen und technischen Eigenschaften der Nachhallplatte. *Rundfunktechnische Mitteilungen* 2:111-116
- Ledergerber S (2002) Application: How to surround. <http://www.vista7.com/e/surround/index.aspx>. Zugriff 10.7.2007
- Lehmann P, Wilkens H (1980) Zusammenhang subjektiver Beurteilung von Konzertsälen mit raumakustischen Kriterien. *Acustica* 45:256-268

- Lund T (2006) Control of Loudness in Digital TV. NAB BEC Proceedings, S 57–65.
- Maempel H-J (2001) Klanggestaltung und Popmusik. Synchron, Heidelberg
- Maempel H-J (2007) Technologie und Transformation. Aspekte des Umgangs mit Musikproduktions- und -übertragungstechnik. In: Motte-Haber H, Hans Neuhoﬀ (Hrsg) Musiksoziologie (Handbuch der systematischen Musikwissenschaft; 4). Laaber, Laaber, S 160–180.
- Martland P (1997) EMI. The first 100 years. **Amadeus Press, London**
- Meyer J (1991) Zur zeitlichen Feinstruktur von Vibrato-Klängen. Bericht der 16. Tonmeistertagung, Karlsruhe, S 338–350
- Mikael L (2005) Hammond & Leslie page. http://captain-foldback.com/main_page.htm. Zugriff 30.3.2007
- Moog Music Inc. (Hrsg) (2003a) Understanding and using your moogerfooger® MF-103 Twelve Stage Phaser. Ashville / NC, <http://www.moogmusic.com/manuals/mf-103.pdf>. Zugriff 28.3.2007
- Moog Music Inc. (Hg) (2003b) Understanding and using your moogerfooger® MF-103 Twelve Stage Phaser. Ashville / NC, <http://www.moogmusic.com/manuals/mf-102.pdf>. Zugriff 28.3.2007
- Moorer JA (1979) About this reverberation business. Computer Music Journal 3/2:13–28
- Muzzulini D (2006) Genealogie der Klangfarbe. Bern et al.: Lang. (Varia musicologica; 5). Zugl. Diss. (2004), Zürich: Univ.
- Motte-Haber H (2005) Modelle der musikalischen Wahrnehmung. Psychophysik – Gestalt – Invarianten – Mustererkennen –
- Neuronale Netze – Sprachmetapher. In: Motte-Haber H, Günther Rötter (Hrsg) Musikpsychologie (Handbuch der systematischen Musikwissenschaft; 3). Laaber, Laaber, S 55–73
- Neoran IM (2000) Surround Sound Mixing using Rotation, Stereo Width, and Distane Pan Pots. 109th AES Convention, Los Angeles, Preprint 5242
- Neumann AC (1998) The Effect of Compression Ratio and Release Time on the Categorical Rating of Sound Quality. J Acoust Soc Amer 103/5:2273–2281
- Nitsche P (1978) Klangfarbe und Schwingungsform (Berliner musikwissenschaftliche Arbeiten; 13). Katznbichler, München, zugl Diss (1974) Techn Univ, Berlin
- Ranum J, Rishøj K (1986) TC 2290 Dynamic Digital Delay + Effects Control Processor Owner's Manual. <http://www.tcelectronic.com/media/d86b1147172e4c87a0fc6a6d3d1dd51c.pdf>. Zugriff 16.2.2007
- Reinecke H-P (1978) Musik im Original und als Reproduktion. In: Roads C (2001) Microsound. Kommunikationstechnik 77. MIT Press, Cambridge / MA, S 140–151
- Rettinger M (1945) Reverberation Chambers for Re-Recording. Journal of the Society of Motion Picture Engineers 46:350–357
- Sandmann T (2003) Effekte & dynamics: Das Salz in der Suppe jeder Musik-Mischung: Technik und Praxis der Effekt- und Dynamikbearbeitung. 3. Aufl, PPV, Bergkirchen
- Schlemm W (1997) Musikproduktion. In: Ludwig Finscher (Hrsg) Die Musik in Geschichte und Gegenwart, Bd 6, 2. Ausg, Bärenreiter, Kassel, Sp.1534–1551
- Schroeder MR (1962) Natural sounding artificial reverberation. J Audio Eng Soc 10/3:219–223
- Schroeder MR (1975) Diffuse sound reflection by maximum-length sequences. J Acoust Soc Amer 57
- Schroeder MR, Logan BF (1961) 'Colorless' Artificial Reverberation. J Audio Eng Soc 9/3: 852–853
- TC electronic (Hrsg) Intonator Vocal Intonation Processor Bedienungsanleitung. http://www.tcelectronic.com/media/Intonator_%20Man_DE.pdf. Zugriff 16.2.2007
- Terhardt E (1986) Gestalt Principles and Music Perception. In: Perception of Complex Auditory Stimuli. Erlbaum, Hillsdale/NJ, S 157–166.
- Theile G (1980) Über die Lokalisation im überlagerten Schallfeld. Diss TU Berlin
- Raffaseder H (2002) Audiodesign. Fachbuchverl Leipzig im Hanser Verl, München
- Ungeheuer E, Supper M (1995) Elektroakustische Musik. In: Ludwig Finscher (Hrsg) Die Musik in Geschichte und Gegenwart, Bd 6, 2. Ausg, Bärenreiter, Kassel, Sp 1717–1765
- Vary P, Martin R (2006) Digital Speech Transmission - Enhancement, Coding & Error Concealment. Chichester: John Wiley and Sons
- Vowinkel A (1995) Collagen im Hörspiel. Die Entwicklung einer radiophonen Kunst. Königshausen & Neumann, Würzburg
- Wagenaars WM, Houtsma AJM, van Lieshout RAJM (1986) „Subjective Evaluation of Dynamic Compression in Music“. J Audio Eng Soc 34(1/2):10–18

Wagner K (1997) Zur Lautheit von Rundfunkprogrammen. Radio, Fernsehen, Elektronik: die Zeitschrift der Unterhaltungselektronik 46(3):42-46.

Wave Arts Inc. (Hrsg) (2005) Panorama User Manual. Arlington/MA. <http://www.wavearts.com/pdfs/PanoramaManual.pdf>. Zugriff 19.7.2007