

Produktionsspezifischer Wandel im Sound von Popmusik

DANIEL SCHOLZ

EINLEITUNG

„What, however, is this objective sound? We cannot entirely take it for granted. It does not exist in any sort of written form but only, ultimately, in our (faulty) memories. All visual representations of it are just that – representations“

(MOORE 2003: 9).

Mit dieser Feststellung verdeutlicht Allan Moore die Problematik, inwieweit eine angemessene und objektive Auseinandersetzung mit dem Parameter Sound in populärer Musik überhaupt möglich ist. Gleichzeitig verweist er damit auf die Diskrepanz zwischen einer auditiv-flüchtigen und einer visuell-statischen Dimension – zwei grundverschiedene Ebenen also, die hinsichtlich der menschlichen Wahrnehmung ihre Entsprechung in den Sinnen *Hören* einerseits und *Sehen* andererseits finden. In der Verschriftlichung wissenschaftlicher Erkenntnisse sind klangliche Phänomene jedoch letztlich immer nur über die visuelle, und somit eigentlich ungeeignete Ebene vermittelbar. Verfolgt man diesen Gedanken noch etwas weiter, so wird klar, dass selbst eine Wellenformdarstellung der originären klanglichen Information nur einen abbildhaften Verweis auf das darstellt, was von geeigneten Wandler-, Verstärker- und Wiedergabe-Systemen in Schall umgewandelt und schließlich gehört werden kann. Allein durch das *Lesen* einer (nicht-elementaren) Wellenform ist es – anders als beim Lesen einer Partitur – nicht möglich, sich eine innere

klangliche Vorstellung zu verschaffen, beispielsweise, welche Instrumente, Töne oder Harmonien erklingen. Umgekehrt ist es ebenso wenig möglich, die über das Ohr aufgenommenen Fluktuationen im Schalldruck in eine analoge, grafische Wellenform zu transkribieren. Und dennoch kommt man im wissenschaftlichen Diskurs nicht umhin, Mittel und Wege der visuellen Repräsentation (beispielsweise durch Wellenformdarstellungen oder Spektralanalysen) zu finden, mittels derer sich adäquate Aussagen über das klanglich Gegebene treffen lassen. Idealerweise sollten solche Aussagen auch nachvollziehbar und überprüfbar sein – handelt es sich dabei doch um zentrale wissenschaftliche Ansprüche.

Das traditionelle Mittel der Notation ist als Repräsentation klanglicher und insbesondere produktionstechnischer Aspekte umso ungeeigneter, je spezifischer und je substanzieller letztere sind. Denn wie könnte man die Kompression einer Gitarre, das Pumpen eines Pads, das Delay einer Stimme oder die räumliche Staffelung der Instrumente im Mix angemessen notieren, wenn nicht durch ergänzend beschreibende Worte? Es dreht sich also letztlich um produktionstechnische Faktoren, die den Sound populärer Musik seit jeher maßgeblich geprägt haben und dies – einhergehend mit technischem Fortschritt und klangästhetischem Wandel – heute in zunehmender Weise tun. Albrecht Schneider verweist außerdem auf den Umstand, dass notierte Tonhöhen im akustischen Sinne und mit einer *singulären Frequenzangabe* lediglich auf einen klangneutralen Grundton verweisen, nicht jedoch auf ein spezifisches Obertonspektrum, durch das der spezifische Klangcharakter eines Instruments aber erst entsteht (Schneider 2002: 108). Aus dieser Perspektive ist die Feststellung von Tilo Hähnel et al. naheliegend, dass es am Beispiel der von ihnen verwendeten und als *Analysetools* bezeichneten Tonhöhenkurven und Spektrogramme „sinnvoll sein kann, auf Darstellungs- und Vermittlungsformen zurückzugreifen, die über die europäische Notenschrift oder eine sprachliche Charakterisierung des Klanggeschehens hinausgehen“ (Hähnel/Marx/ Pfleiderer 2014: 19). Da bei Musik, die in erster Instanz produziert und nicht notiert wurde, letztlich nur das per Tonträger übermittelte Audiosignal zur Verfügung steht, so kann die analoge Visualisierung als dessen Wellenform als die unmittelbarste Repräsentation des Gehörten – des eigentlichen Primärtexts im Moore’schen Sinne (vgl. Moore 2001: 35) – angesehen werden. Als Repräsentanten von Schalldruckschwankungen und als Träger nicht musikalischer, sondern akustischer Strukturen, werden Audiosignale von Rolf Großmann auch als „Materialität zweiter Ordnung“ bezeichnet, während deren digital kodierte Repräsentation mitsamt der untrennbar zugehörigen Kon-

figuration der für die Rückführung notwendigen Apparate (Sampler, Computer, Interfaces etc.) gar als „Materialität dritter Ordnung“ angesehen wird (vgl. Großmann 2004: 97 ff.). Dazu sollte angemerkt werden, dass Audiosignale als kodierte Klanginformation, egal ob analog oder digital, letztlich dennoch jegliche musikalische Struktur in sich bergen.

Die bereits existierenden Ansätze, Sound-Analysen auf Basis des Audiosignals zu erstellen, lassen sich grundsätzlich einteilen in solche, die zur groben Veranschaulichung lautmalerischer Klangaspekte (z.B. offensichtliche Treppen-Dynamik oder Crescendi) lediglich auf die Originalwellenform verweisen, und solche, die das Originalsignal zunächst einem technischen Analyseprozess unterwerfen, dessen Ergebnis typischerweise in Form eines Sonogramms oder Spektrogramms präsentiert wird. Anhand von Sonogrammen lassen sich Einblicke in die Frequenzverteilung innerhalb eines zeitlichen Ausschnitts gewinnen, jedoch haben sie hinsichtlich ihrer Aussagekraft eine inhärente Schwäche: Da die Auflösungen in Frequenz- und Zeit-Domäne in antiproportionaler Abhängigkeit stehen, bedeutet eine feinere frequenzielle Auflösung eine umso unschärfere zeitliche – und umgekehrt. Außerdem ist die Gesamterscheinung meist relativ verwaschen und unleserlich und in der jeweiligen frequenziellen Staffelung nur schwer zu interpretieren; von nachvollziehbaren, semantischen Konnotationen einmal ganz zu schweigen. Visualisiert wurden mit Sonogrammen jedenfalls unter anderem Bass-Drum und HiHat in elektronischer Tanzmusik (vgl. z.B. Zeiner-Henriksen 2010) oder auch Oberton-Spektren einer E-Gitarre (vgl. z.B. Schneider 2002: 111 ff. oder Berger und Fales 2005). Von besonderem Interesse ist in der Forschung insbesondere der Klang der menschlichen Stimme als eines der zentralen Elemente von Popmusik. Erste Versuche, den klanglichen Detailreichtum einer Singstimme durch Spektralanalysen zu untersuchen, wurden bereits in den 1930er Jahren unternommen. So berichtete Wilmer Bartholomew 1934 über die Sängerformanten – ausgeprägte Energiezentren oder Resonanzspitzen zwischen 2,5 kHz und 3,3 kHz (Bartholomew 1934). Dieser grundlegende Ansatz, eine allgemeine Stimmfarbe anhand der Oberton-Charakteristik sichtbar zu machen, wurde in den folgenden Jahrzehnten immer wieder aufgegriffen (vgl. exemplarisch Winckel 1953, Sundberg 1968, Bunge 1977, Fördermayr 1985, Fördermayr/ Deutsch 1989). Es wurde sogar der Versuch unternommen, das *Knarzen* der Stimme in Sonogrammen wiederzufinden (Lacasse 2010). Und jüngst präsentierten Tilo Hähnel, Tobias Marx und Martin Pfeleiderer (2014) einige *Methoden zur Analyse der vokalen Gestaltung populärer*

Musik. Anhand von Tonhöhenkurven visualisieren sie – über die in ihren Augen reduktionistisch-problematische Notentranskription hinausgehend – Details des Tonhöhengleitens, der Vibratogestaltung und der Vokalartikulation.

Ganz in Großmanns Sinne dürften all solche teils algorithmisch gestützten, auf den Parameter Sound abzielenden Material-Analysen – vergleichbar mit dem DJ-Set oder Sampler – letztlich auch als eine „spezielle Praxis der kulturellen Aneignung elektronischer Medien“ verstanden werden, geht es doch auch hier um einen „gestalterischen Gebrauch“ der technischen Signale (vgl. Großmann 2004: 91). Durch diese Aneignungsformen findet zwar nicht unbedingt eine tatsächlich erklingende, ästhetische Neuproduktion wie bei einer auf Samples basierenden DJ-Performance statt, dennoch wird durch sie eine „metaphorische[n] Übertragung technischer Modelle und Begriffe in den Diskurs“ (ebd.: 95) ermöglicht. Auf Basis dieser Vorüberlegungen und mit einem Fokus auf stereofone Klangaspekte werden im Folgenden einige ältere Stücke und darauf bereits angewendete Analysemethoden einigen neueren Stücken gegenübergestellt, um detailliertere Erkenntnisse hinsichtlich produktions-technischer und klanggestalterischer Paradigmenwechsel zu gewinnen. Dabei ist es das Ziel, anknüpfend an die genannten, üblichen Visualisierungsformen, andere methodische Zugänge oder technische Aneignungsformen zu erproben, um möglichst spezifische Einblicke zu erhalten. Abschließend wird unter anderem der Frage nachgegangen, inwieweit eine der etablierten Analysemethoden noch geeignet ist, ein aktuelleres Klanggeschehen adäquat zu erfassen oder vielmehr auf eine Epoche der 1960er und frühen 70er Jahre beschränkt ist.

DER MIX IM RÄUMLICHEN KONTEXT

Die Stereophonie stellt mit allen daran geknüpften räumlichen Aspekten einen zentralen Parameter klanglicher Gestaltung dar, der offenbar einer lediglich auf Höreindrücken basierenden Analyse am ehesten zugänglich ist. So wurden in diesem Zusammenhang Konzepte und Begriffe geprägt, wie z.B. die „Sound Stage“ (Moylan 1992: 48) oder das „Functional Staging“ (Zagorski-Thomas 2010). Insbesondere der *Space*-Begriff erhielt eine größere Aufmerksamkeit im fachlichen Diskurs (vgl. Moylan 2009, Camilleri 2010, Michelsen 2012, Holland 2013). Bei dem verbreiteten *Sound-Box*-Konzept von Allan Moore (Moore 1992) geht es im Wesentlichen darum, die durch Höranalyse gewonnenen, räumlich-stereofonen Erkenntnisse durch Piktogramme der einzelnen Instrumente geeignet in einem transparenten Quader anzuordnen. Auch

heute noch scheint die Sound-Box eine bevorzugte Methode zu repräsentieren, wenn es darum geht, die Staffelung der Komponenten einer Mischung im räumlichen Kontext zu visualisieren (vgl. z.B. Braae 2015, Figure 1). In Zusammenarbeit mit Ruth Dockwray vermag Moore mit der Untersuchung verschiedener Stücke im Zeitraum von 1965-1972 schließlich eine Kategorisierung verschiedener Panning-Paradigmen zu präsentieren, die als *Cluster*, *Triangular*, *Diagonal* und *Dynamic* definiert werden. Den Autoren zufolge sei ab 1972 die Etablierung einer *normativen* Positionierung von Klangquellen innerhalb der Sound-Box vollzogen, was dem *diagonalen Mix* entspräche (vgl. Dockwray/ Moore 2010). Inwieweit die Sound-Box-Analysemethod jedoch geeignet ist, das als *diagonal* oder *normativ* bezeichnete Panning-Paradigma von Produktionen ab 1972, oder gar aktueller Produktionen adäquat abzubilden, wird im Artikel selbst nicht beleuchtet.

DIE M/S-KODIERUNG

Stereo-Signale werden üblicherweise durch einen linken und einen rechten Kanal repräsentiert. Darüber hinaus gibt es jedoch auch die Repräsentation durch eine *Mitte* und eine dazu orthogonale *Seite*. Anhand einer solchen M/S-Aufspaltung lässt sich aus analytischer Perspektive ermitteln, welche Elemente des Mixes entweder mittig oder seitig platziert sind oder wie stark ausgeprägt deren jeweilige Stereobasisbreite (Breitenwirkung), das heißt Amplitude im Seitensignal, ist. Liegt beispielsweise ein Mono-Element exakt mittig (z.B. Bass Drum, Snare Drum, Stimme etc.), so fehlt dieses Element vollständig im Seitensignal. Diese Technik, zunächst gewinnbringend eingesetzt bei der Stereoschallplatte, hielt aber auch Einzug in den Rundfunk und ist heute nach wie vor sowohl in der Mikrofonierungstechnik als auch im Mastering ein unverzichtbares Mittel. All diese Anwendungen haben die Gemeinsamkeit, dass unter Verzicht des Seitensignals – sei es aus Gründen schlechter Übertragung (Rundfunk) oder aufgrund von Limitierungen des Abspielgeräts (Mono-Plattenspieler oder Mono-Abhörsystem) mit dem Mittensignal immerhin genau dasjenige Signal erhalten bleibt, das im Prä-Stereo-Sinne die optimale Mono-Fassung repräsentiert. Insofern kann die Mitte als unverzichtbare Basis angesehen werden, während die Seite eine – notfalls entbehrliche – qualitative Bereicherung darstellt (vgl. Abschnitt *Mono-Kompatibilität*, s.u.).

Um eine M/S-Aufteilung vorzunehmen, mischt man den linken und den rechten Kanal zu jeweils 50%, oder: $M = (L+R)/2$, und erhält damit das Mittensignal. Das Seitensignal gewinnt man hingegen durch die Differenz von linkem und rechtem Kanal, wiederum zur Hälfte der jeweiligen Lautstärke: $S = (L - R)/2$. Indem man das abzuziehende Signal gegenphasig, oder phaseninvertiert, zum anderen Signal hinzumischt, entsteht diese Differenz. Die Notwendigkeit, jeweils durch 2 zu teilen, ergibt sich lediglich aus Gründen der Lautstärkekompensation, denn sonst würde die Summe zweier maximierter Signale L und R zu ca. 6 dB Clipping (Übersteuerung), also ungefähr einer Verdoppelung der Lautstärke führen. Auf Basis dieser Zerlegung des Stereosignals in die Komponenten Mitte und Seite lässt sich schließlich das ursprüngliche Stereosignal verlustfrei zurückgewinnen, oder rückmatrizieren: $L = M + S$ sowie $R = M - S$. Diese Technik ist übrigens auch als Aufnahmeverfahren mit Hilfe zweier Mikrofone geeignet (vgl. Dickreiter u. a. 2014: 253 ff.): Das Mittensignal wird durch eine nach vorn ausgerichtete Kugelcharakteristik eingefangen, während eine seitlich ausgerichtete Achtercharakteristik das Seitensignal aufnimmt. Unterwirft man diese beiden Signale der M/S-Rückmatrizierung, so erhält man daraus das vertraute L/R-Stereo-Signal. Wie sich im weiteren Verlauf herausstellen wird, sind die Signale M und insbesondere S imstande Einblicke zu liefern, die aus der üblichen L/R-Aufteilung schlichtweg nicht gewonnen werden können. Bei dieser besonderen Art der stereofonen Kodierung handelt es sich um ein mächtiges Werkzeug nicht nur für spezielle Verfahren im Prozess des Mixing und/oder Mastering, sondern in einzigartiger Weise auch für die Sound-Analyse, wie im Folgenden gezeigt wird.

STEREO-PARADIGMEN

Um an ein älteres Stück anzuknüpfen, das mit der oben beschriebenen Sound-Box-Methode analysiert wurde, möge der 1967er-Titel „There Goes My Everything“ von Engelbert Humperdinck, produziert von Peter Sullivan, betrachtet werden: Die Sound-Box-Analyse dieses Stücks zeigt Drums, Bass, Gitarre links, Streich- und Blasinstrumente rechts, Vocals mittig (vgl. Dockwray/ Moore 2010: 188 f.) – also ein relativ freies Panning der Instrumente im Stereofeld, das bisweilen auch als *Knüppelstereofonie* bezeichnet wird. Die in ihrer Abbildung veranschaulichten konträren räumlichen Positionierungen der verschiedenen Instrumente spiegeln sich auch ganz klar als deutlich sicht-

bare Unterschiede zwischen linkem und rechtem Kanal in der Wellenformdarstellung der Originaldatei wider, siehe Abb. 1 (oben) im Ausschnitt 1:33.6-1:40.5: Während z. B. im hinteren Abschnitt links einige Ausschläge durch die E-Gitarre, Klavier und die mit Besen gespielten Drums zu verzeichnen sind, liegt im rechten Kanal davon unberührt der eher flächige Streicherteppich. Da die besagten Elemente ganz nach links gemischt wurden, sind deren Ausschläge im rechten Kanal nicht vorhanden. Das bedeutet umgekehrt, dass sich diese Elemente sowohl im Mitten- als auch im Seitensignal wiederfinden, vgl. Abb. 1 (unten).

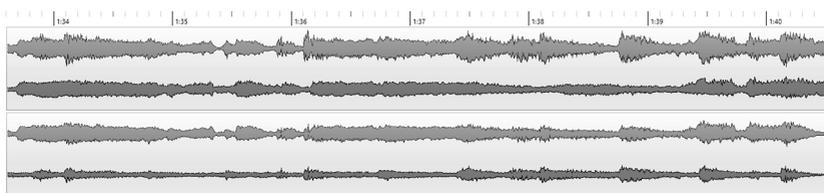


Abb. 1 Engelbert Humperdinck (1967): „There Goes My Everything“ (1:33.6-1:40.5), Wellenformdarstellung des L/R-Originalsignals (o.) und der M/S-Aufteilung (u.).

Als ein exemplarischer Vertreter einer offensichtlich moderneren Klangästhetik sei nun ein Mix von Chris Lord-Alge von 2004 den bisherigen Erkenntnissen gegenübergestellt: Green Day – „American Idiot“, produziert von Rob Cavallo und Green Day.

Die Abb. 2 zeigt einen kurzen Ausschnitt aus dem Chorus, und zwar wiederum oben in der originalen L/R-Darstellung, unten in der M/S-Aufteilung. Anhand der oberen Wellenformdarstellung lässt sich feststellen, dass es keine offensichtlichen Unterschiede zwischen dem linken und dem rechten Kanal gibt. Dies bedeutet zum einen, dass alle Transienten der Drums links genauso stark ausgeprägt sind wie rechts. Zum anderen sind die Pegel links und rechts grundsätzlich gleich laut; Amplitudenverläufe, Verdichtungen, Einschnitte etc. sind im Wesentlichen identisch. Diese hohe Ähnlichkeit beider Kanäle, die also in diametralem Gegensatz zu den Beispielen der Sound-Box-Analysen aus den späten 60ern und frühen 70ern steht, hat natürlich zur Folge, dass sich die grundlegende Wellengestalt von L und R ebenso im Mittensignal niederschlägt. Im ersten Drittel des Wellenform-Ausschnitts ist übrigens noch der Gesang dabei, der ab dem zweiten Drittel aussetzt. Während man die nochmals gesteigerte Klangdichte des ersten Drittels gegenüber den hinteren zwei

Dritteln in L und R, noch besser sogar in M nachvollziehen kann, ist das S-Signal hiervon völlig unberührt, d. h. die Stimme ist dort nicht enthalten. Ebenso fehlen die Drums fast vollständig im Seitensignal. Dies bezeugt, dass primär die mittig gemischten Mono-Mikrofone (close mics) von Bass Drum und Snare Drum und nur ein geringer Signalanteil der obligatorischen Stereo-Overhead- oder Stereo-Raum-Mikrofonierungen verwendet wurden. Auch der E-Bass fehlt vollständig im Seitensignal, so dass auch hier von einer mittig gemischten Mono-Aufnahme des Basses ausgegangen werden muss. Das einzige, das letztlich im Seitensignal übrigbleibt, ist eine fulminante Wand aus E-Gitarren. Auf welche Weise möglich gemacht wird, aus der an sich *punktförmigen* Schallquelle einer Gitarre ein derart opulentes Stereo-Signal zu generieren, wird weiter unten im Abschnitt Pseudostereofonie und Doppelungs-Technik untersucht.

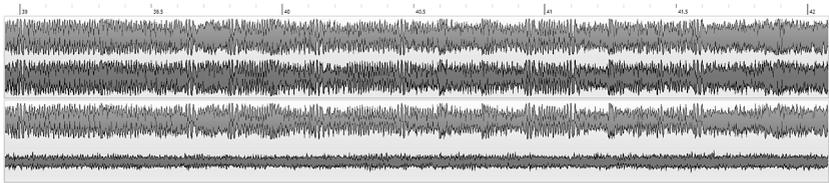


Abb. 2 Green Day (2004): „American Idiot“ (38.9-42.1), Wellenformdarstellung von L/R-Originalsignal (o.) und M/S-Aufteilung (u.).

Aus Abb. 2 und Abb. 1 wird jeweils von oben nach unten ersichtlich: Würde man nur den linken Kanal, nur den rechten Kanal, oder nur das Mittensignal des Titels „American Idiot“ (Green Day 2004) abhören, so erhielte man keine signifikanten Abweichungen im Misch-Verhältnis aller Instrumente – denn die Signale L, R und M sind hinsichtlich Transienten und Amplitudenverläufen äußerst ähnlich. Hört man sich hingegen den linken und rechten Kanal des Titels „There Goes My Everything“ (Humperdinck 1967) jeweils einzeln und im direkten A/B-Vergleich an, so fällt auf, dass sich die Präsenz der einzelnen Instrumente derart substanziiell unterscheidet, dass man eigentlich schon von zwei grundverschiedenen Mischungen oder alternativen Editionen sprechen müsste, von denen eine nach links und die andere nach rechts gelegt worden ist; der rechte Kanal ist, im Vergleich zum linken, praktisch ein alternativer *Remix* des Songs; das Mittensignal ist gewissermaßen ein *Kompromiss* aus beiden Fassungen und stellt somit nur bedingt eine optimale, dezidierte Mono-Fassung dar. Im frühen Mixing-Paradigma musste man sich offenbar stets entscheiden: Will man einen fortschrittlichen, aber in der Wirkung womöglich

diffus-zerstreuten (aus heutiger Sicht eher effekthascherischen) Stereo-Mix erstellen, oder bleibt man dem Ideal des fokussierten, guten alten Mono-Mixes treu? Wohl nicht ohne Grund waren Produzenten wie Phil Spector oder Brian Wilson trotz der aufkeimenden Stereo-Technologie überwiegend der Meinung, ihre klangliche Vision besser über einen dezidierten Mono-Mix zum Ausdruck bringen zu können – man denke nur an Alben wie *Love Is All We Have to Give* (The Checkmates Ltd. 1969) oder *Pet Sounds* (Beach Boys 1966). Jedoch die Vorzüge beider Domänen in einem Mix zu verkörpern, schien schlichtweg nicht möglich. Ob und inwieweit dies heute anders ist, wird am Beispiel „American Idiot“ (Green Day 2004) nachfolgend erörtert.

MONO-KOMPATIBILITÄT

Dieser exemplarische Mix von Chris Lord-Alge vermag einen grundlegenden Trend anzuzeigen, der sich, wie Bartlett bereits vermutete, für eine zeitgemäße Klangästhetik als *typisch* charakterisieren lässt: Während man die Elemente Bass, Snare Drum, Bass Drum und Vocals in der Mitte platziert, werden Keyboards und Gitarren nach links und rechts verlagert (vgl. Bartlett und Bartlett 2002: 289). Dadurch, dass diese vier dominierenden Elemente monofon und mittig, also im wahrsten Wortsinn *zentral* positioniert sind, fehlen sie mehr oder weniger vollständig im Seitensignal; alle *flächigen* Elemente, wie z. B. eine Gitarrenwand, Keyboards, Synthesizer, Streicher etc. mit stark ausgeprägtem Seitenanteil, sind hingegen nur schwach in der Mitte, dafür umso stärker im Seitensignal angesiedelt und sorgen in klanglicher Hinsicht für eine zusätzliche Dimension und Klangfülle. Insofern trägt das moderne Mixing-Paradigma erstaunlicherweise einem fast 60 Jahre alten Prinzip Rechnung: der *Mono-Kompatibilität*. Diese ursprünglich technische Erfordernis sollte gewährleisten, dass die neuen Stereo-Technologien immer noch mit den alten Mono-Geräten funktionierten und somit abwärtskompatibel blieben. So konnte beispielsweise eine Stereo-Schallplatte immer noch auf einem Mono-Plattenspieler abgespielt werden, da hier mit der M/S-Kodierung das ursprüngliche Mono-Signal nun in Form des Mittensignals erhalten blieb; das hinzugekommene Seitensignal wurde in einer 45°-Flanke der Rille eingepreßt, das die alten Plattenspieler schlicht ignorierten. Auch im Rundfunk konnte die Mono-Kompatibilität sichergestellt werden, indem das Mittensignal an die Stelle des bisherigen Mono-Signals trat und das Seitensignal auf eine höhere Trägerfrequenz moduliert und somit von älteren Geräten einfach

nicht verwertet wurde. Demnach besteht eine *qualitative Diskrepanz* zwischen den Komponenten Mitte und Seite: Während in einer L/R-Kodierung beide Kanäle grundsätzlich den gleichen Stellenwert haben, so repräsentiert in der M/S-Kodierung das Mittensignal die wichtigere der beiden Komponenten; als Basis ist es unentbehrlich, während mit dem Seitensignal eine zusätzliche Qualität oder Dimension eröffnet werden kann, aber nicht muss.

Ganz im Sinne dieser technischen Perspektive erfährt die Mono-Kompatibilität nun auch im modernen Mixing-Paradigma eine Entsprechung: Die für den klanglichen Gesamteindruck essenziellen Instrumente werden in der Mitte angesiedelt und transportieren bereits alle wesentlichen klanglichen Informationen. Insofern ist diese Mixing-Ästhetik in optimaler Weise mono-kompatibel. Da im Beispiel „American Idiot“ das Seitensignal praktisch nur aus der Gitarrenwand besteht, könnte man auch sagen, dass es sich – selbst im lange etablierten Stereo-Zeitalter – im Kern um eine solide Mono-Mischung handelt, die um eine Dimension der *Breite* lediglich ergänzt wurde. Dass sich solche Mischungen unter jeder denkbaren Abhörsituation optimal durchsetzen, liegt letztlich daran, dass die zentralen Elemente nicht aus irgendeiner fernen Richtung ertönen (vgl. mit der Knüppelstereofonie, wo es keine stabile Mitte gibt), sondern mit voller Energie aus beiden Lautsprechern gleichermaßen. Somit stellt diese als modern anzusiedelnde Mixing-Ästhetik in ihrem vermeintlichen Fortschritt streng genommen auch einen Rückschritt dar, nämlich in ihrer Priorisierung einer soliden Mono-Fassung – so, als ob die ehemalige, heute praktisch überwundene Maßgabe der Mono-Kompatibilität eine noch stärkere Rolle spielen würde als früher, obwohl sie technisch längst nicht mehr erforderlich ist: denn es gibt keine Mono-Plattenspieler, keinen Mono-Rundfunk und, abgesehen von wenigen Ausnahmen, praktisch auch keine Mono-Wiedergabesysteme mehr. Dennoch wird offenbar so stark mono-kompatibel gemischt, dass eigentlich die heutigen Produktionen viel besser auf den alten Mono-Plattenspielern klängen als die damaligen, in ein Mono-Korsett gezwängten Knüppelstereo-Produktionen. Da es sich – im Gegensatz zu früher – bei der Mono-Kompatibilität also nicht mehr um ein zwingendes technisches Prinzip handelt, schlägt es sich heute eigentlich nur noch als *klangästhetisches Ideal* nieder. Und ein Mix, der nach diesem Ideal gestaltet wurde, scheint zwei Gegensätze in sich zu vereinen: er ist optimaler Mono-Mix und optimaler Stereo-Mix gleichzeitig. Keine der früheren Produktionstechniken und Klangästhetiken hätte dies zu leisten vermocht.

PSEUDOSTEREOFONIE UND DOPPELUNGS-TECHNIK

Ob sich auch das im Rock/Pop unverzichtbare Instrument A-Gitarre dazu eignet, ebenso wie die E-Gitarre als primär im Seitensignal angesiedelte, Fülle erzeugende *Klangfläche* inszeniert zu werden, sei wiederum anhand eines älteren und jüngeren Beispiels untersucht. Betrachten wir mit dem 1970er-Song „Lola“ der Band The Kinks zunächst einen weiteren Titel aus der frühen Zeit der Stereophonie, der bereits mit der Sound-Box-Methode analysiert worden ist; die von Dockwray/ Moore angefertigte Abbildung stellt ein Drumset, eine E-Gitarre, eine A-Gitarre, einen E-Bass und Vocals mehr oder weniger mittig dar (vgl. Dockwray/ Moore 2010: 193). Und tatsächlich zeigt eine Interpretation des Mitten- und Seiten-Signals, dass der gesamte Mix von Beginn an bis Sekunde 34.8 quasi monofon ist, da das Seitensignal in diesem Abschnitt annähernd stumm ist (vgl. Abb. 3).

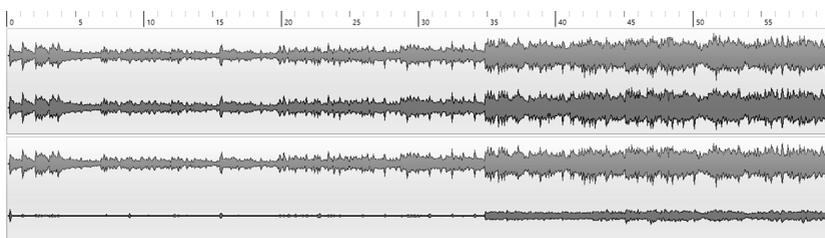


Abb. 3 The Kinks (1970): „Lola“ (0.0-1:0.0), Wellenformdarstellung von L/R-Originalsignal (o.) und M/S-Aufteilung (u.).

Im perfekten Mono müsste es zwar vollkommen stumm sein, aber man kann hier – im Gegensatz zum heutigen Digitalzeitalter – von einigem Übersprechen zwischen Mischpultkanälen, Bandmaschinenspuren und/oder Schallplattenrillenflanken, also von technischen Unzulänglichkeiten alter Analogtechnik, ausgehen. Somit ist auch die A-Gitarre, die von Beginn an mitspielt, offensichtlich monofon aufgenommen und in die Mitte gelegt worden; erst ab Sekunde 34.8 gesellt sich links außen das Klavier, im weiteren Verlauf eine Solo-E-Gitarre als Gegengewicht im rechten Kanal hinzu. Von einer die gesamte Stereobasisbreite ausfüllenden *Klangwand* kann hier aufgrund der nur punktförmig verteilten, von der Mitte abweichenden Instrumente Klavier einerseits und E-Gitarre andererseits, keine Rede sein. Bei diesem Beispiel, das Dockwray und Moore als frühen Vertreter des ab 1972 etablierten *normativen Mixes* erwähnen, wird ein ausgewogenes Stereobild vielmehr dadurch zu erreichen versucht, indem auf eine ausgewogene Balance zwischen jeweils nach

links oder rechts gepannten Instrumenten Wert gelegt wird. Dieser Ansatz kann somit gewissermaßen als ein *Hybrid* zwischen einem Mono- und Knüppelstereo-Mix angesehen werden.

Hören wir in Gegenüberstellung den Beginn des Songs „Over You“ (2006) der amerikanischen Band Daughtry, produziert von Howard Benson: Dieser Track startet ebenfalls mit A-Gitarre und Gesang, allerdings wirkt hier der Gesang von Beginn an geradezu *umhüllt* von einer sehr füllig und breit inszenierten A-Gitarren-Wand. Wie ist es möglich, dass eine an sich relativ monofon-punktförmige Schallquelle derart ausgeprägt stereofon klingen kann? Eine Möglichkeit wäre, die Gitarre mit zwei Mikrofonen aufzunehmen. Falls man über die Wellenform herausfinden könnte, ob die Saiten-Anschläge links und rechts zeitlich identisch sind, so würde immerhin die Intensitätsstereofonie in Frage kommen, wengleich diese kaum geeignet scheint, eine solche Breitenwirkung zu erzeugen. Sähe man stattdessen einen mehr oder weniger konstanten zeitlichen Abstand in den Saitenanschlägen zwischen den beiden Kanälen, so würde dies für eine Mikrofonierung mit Laufzeitstereofonie sprechen. Denkbar wäre allerdings noch ein weiteres Verfahren, das sogar eine monofone Aufnahme der A-Gitarre in ein mehr oder weniger überzeugendes Stereoformat umwandeln könnte: die *Pseudostereofonie*. Unter diesem Begriff ist eine Vielzahl an Möglichkeiten subsumiert, mit denen ein künstliches Stereosignal aus einem Monosignal gewonnen werden kann. Bereits 1977 stellt Ingo Voß einige dieser Verfahren vor, mit denen eine „*scheinbar* plastische Wiedergabe einer Monoaufnahme“ generiert werden kann, „ohne jedoch echte stereofonische Information zu besitzen“ (Voß 1977: 245, Hervorh. i.O.). Etwas jüngeren Darstellungen zufolge wird angeregt, aus dem Mono-Signal ein künstliches – konkret: hochpassgefiltertes und verzögertes – Seitensignal zu konstruieren, mithilfe dessen nach der M/S-Rückmatrizierung ein Pseudo-Stereo-Signal entsteht (vgl. Robjohns 2003). Gemäß Paul White könne auch der linke oder rechte Kanal des Mono-Signals mit einem subtilen Chorus-Effekt versehen werden, was gut bei Pads und Gitarren, weniger gut jedoch bei Vocals und Drums funktioniere. Alternativ könne man einen grafischen Equalizer mit experimentellen Einstellungen zur Gewinnung eines phasen-manipulierten Signals verwenden, das zum mittigen Originalsignal links normal und rechts phaseninvertiert hinzugemischt wird. Auf diese vollständig mono-kompatible Art würden die in Stereo-Signalen üblichen Kammfilter-Effekte simuliert (vgl. White 2000).

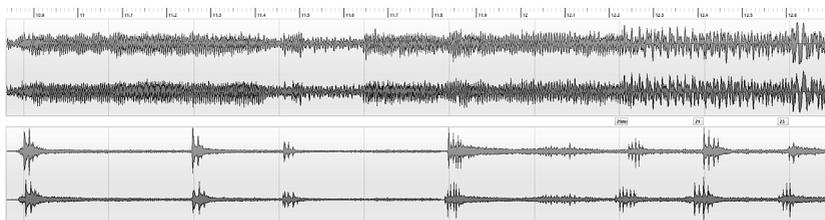


Abb. 4 Daughtry (2006): „Over You“ (10.9-12.7), oben: L/R-Originalsignal; unten: Isolierung von L/R-Doppelung der A-Gitarre per HP-Filter.

Um nun die Antwort auf die aufgeworfene Frage zu finden, mit welchem der erörterten Mittel die ausgeprägte Stereo-Wirkung der A-Gitarre im Song „Over You“ (Daughtry 2006) realisiert wurde, wird im Folgenden eine weitere Methode zur Sound-Analyse vorgestellt, mit der es in diesem Fall – wie gewünscht – möglich werden wird, Einblicke in die zeitlichen Positionen der einzelnen Saitenanschläge zu erhalten: die *Hochpassfilterung*. Wendet man in diesem Fall einen Hochpassfilter bei 15 kHz auf das originale L/R-Audiosignal an und erhöht noch die Lautstärke aus Gründen der besseren Sichtbarkeit, so kommen alle hochfrequenten Transienten zum Vorschein (siehe Abb. 4). Bei jeder ausgeprägten, sehr kurzen Pegelspitze handelt es sich tatsächlich um den Anschlag des Plektrums an einer der sechs Gitarrensaiten. Gitarrist_innen unterscheiden zwar zwischen schnell geschlagenen und gezielt langsam arpeggierten Strums, jedoch wird durch die Grafik offensichtlich, dass selbst die schnell ausgeführten Strums bei hinreichend hoher zeitlicher Auflösung letztlich Arpeggien sind, wenngleich äußerst schnelle. Die Zeit, die jeweils zwischen erstem und letztem Saitenanschlag vergeht, konnte hier mit durchschnittlich ca. 34 ms ermittelt werden – schnell genug offenbar, um dem menschlichen Ohr den Eindruck eines einzelnen Anschlags zu suggerieren.¹ Die keiner Regel folgenden, feinen zeitlichen Unterschiede zwischen den Saitenanschlägen im linken und rechten Kanal belegen nun, dass keine der oben beschriebenen und in Erwägung gezogenen Methoden der Stereo-Erzeugung zum Einsatz kam: Weder sind die Saitenanschläge links und rechts zeitlich

1 Der maximale zeitliche Abstand zweier Klangereignisse, bei dem man sich diese noch als ein einziges Ereignis zurechthören kann („temporal integration“), ist von Eddins/ Green (1995: 237) mit 100-200 ms ermittelt worden; dagegen liegt die Hörschärfe zur tatsächlichen Identifizierung der zwei Ereignisse („temporal acuity“) mit einem Bruchteil einer Millisekunde bis zu 30 ms deutlich niedriger (ebd.).

identisch (Intensitätsstereofonie) noch zeitlich konstant versetzt (Laufzeitstereofonie); ebenso wenig lässt sich eine sonstige Regularität erkennen, die sich einem der dargelegten Pseudostereofonie-Verfahren zuordnen ließe. Stattdessen handelt es sich zweifelsfrei um zwei voneinander unabhängige, das heißt nacheinander durchgeführte Mono-Aufnahmen (Takes), die durch entgegengesetztes Panning zu gefühlt *einer* virtuellen, breiten Stereo-Gitarre verschmelzen. Dieses Verfahren wird auch als Doppelung, *Doubling* oder *Double Tracking* bezeichnet; das damit gewonnene Signal enthält aufgrund subtiler Fluktuationen hinsichtlich Intonation, Klangfarbe, Timing, Lautstärke und Anschlagstärke eine ausgeprägte Stereo-Wirkung und liefert somit ein ebenso ausgeprägtes Seitensignal. Zur Erinnerung: Im Seitensignal schlagen sich ebendiese Differenzen zwischen den beiden Kanälen, in diesem Fall also zwischen den beiden Takes, nieder. Die Methode der Links/Rechts-Doppelung kann eigentlich als ein Zwitterwesen zwischen echter Stereofonie und der Pseudostereofonie angesehen werden, wenngleich erstere nicht unbedingt intensiver klingen muss. Im Gegenteil: Mit der Double Tracking Methode ist es – im Gegensatz zu einer stereofon mikrofonierten Aufnahme nur eines Takes – immer leicht, eine maximale Stereobasisbreite zu erzielen; und diese wird sogar noch umso intensiver, je inkonsistenter der/die Musiker_in die beiden Takes nacheinander einspielt. Denn dadurch vergrößern sich bis zu einem gewissen Grad die für ein ausgeprägtes Seitensignal erforderlichen Differenzen zwischen den beiden Kanälen.

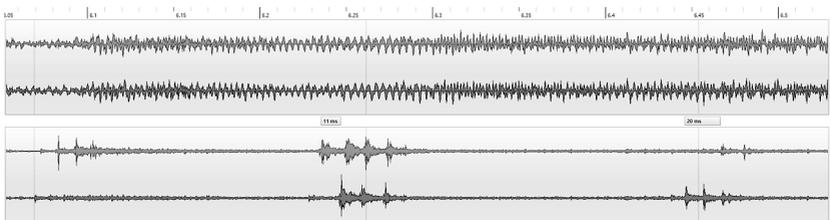


Abb. 5 Daughtry (2006): „Over You“ (6.1-6.5), L/R-Originalsignal und L/R-Doppelung der A-Gitarre per HP-Filter – vergrößerte Darstellung dreier Strums.

Auffällig ist bei der Betrachtung des hochpassgefilterten Signals noch, dass die Abweichungen im Timing der beiden Takes hinten größer sind als vorne (siehe Abb. 4). Dem heute üblichen Editing-Perfektionismus entsprechend ist davon auszugehen, dass zeitlich zu stark abweichende Takes – zumal in exponiertem Kontext wie der hier schlank besetzten ersten Strophe – geschnitten

und geschoben wurden, so dass keine zu großen zeitlichen Abweichungen für Irritation sorgen. Denn schließlich muss man immer berücksichtigen, dass das menschliche Ohr sehr empfindlich auf L/R-Laufzeitunterschiede reagiert – zeichnet diese Eigenschaft doch letztlich entscheidend unsere Fähigkeit zum räumlichen Hören aus. Zu große Timing-Abweichungen zwischen den Kanälen würden zu einem wahllosen Springen der wahrgenommenen Position der Schallquelle im Stereofeld führen. Dennoch wird dies bei den letzten drei Strums offenbar in Kauf genommen; die Laufzeitunterschiede der beiden Takes bzw. Kanäle liegen bei 29 ms, 21 ms und 23 ms – Werte also, die immerhin noch innerhalb der durchschnittlichen Strum-Dauer von 34 ms (s. o.) liegen. Für eine ausgeprägte L/R-Ortung basierend auf Laufzeitunterschieden sind diese Werte jedoch bereits groß genug; man wird diese drei Strums immer rechts-dominant hören, obwohl die Lautstärken gar nicht so unterschiedlich sind. Dass hier nicht editiert wurde, liegt möglicherweise darin begründet, dass diese Strums bereits im Auftakt zur zweiten Hälfte der Strophe liegen, wo der einsetzende Bass die Aufmerksamkeit des Hörers absorbiert und direkt danach die Gitarre in ihrer Dominanz durch das einsetzende Schlagzeug etwas in den Hintergrund rückt. In einem kürzeren, dafür vergrößerten Ausschnitt der Sekunde 6.1-6.5 (siehe Abb. 5) wird sogar ersichtlich, dass die beiden Takes nicht nur hinsichtlich Anschlagstärke und Timing variieren, sondern teilweise auch in der Anzahl der mit dem Plektrum *erwischten* Saiten. Dass es sich in den meisten Fällen um weniger als die sechs vorhandenen Saiten handelt, ist angesichts der Tatsache verständlich, dass zum einen einige Akkordgriffe nur aus fünf oder vier angespielten Saiten bestehen; zum anderen ist die Schlagbewegung beispielsweise bei leichten Achteln deutlich beiläufiger und wird mit weniger Druck ausgeführt, um eine gute dynamische Abstufung im Pattern zu erzielen. Zwei Messpunkte im zeitlichen Versatz zwischen linkem und rechtem Take wurden mit 11 ms bzw. 20 ms ermittelt; auch diese Werte liegen im Rahmen der durchschnittlichen Strum-Dauer von 34 ms (s. o.).

Bei weniger transientenreichen Signalen ist das L/R-Doppelungs-Verfahren generell nicht so dezidiert, dafür teilweise aber großflächiger visualisierbar. Anhand des Stücks „Dirty Window“ von Metallica (2003), produziert von Bob Rock und Metallica, lässt sich im Ausschnitt 13.1-32.0 Sekunden – ohne ein technisches Verfahren zu bemühen – unmittelbar an der originalen Wellenform ablesen, dass im Intro nur die linke E-Gitarre spielt, während die rechte aussetzt (vgl. Abb. 6); bei den kurzen Einwüfen handelt es sich um Drum Fill-ins.

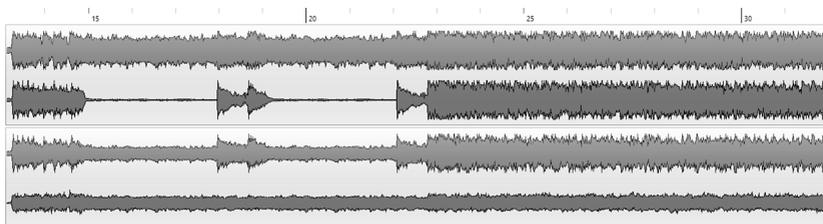


Abb. 6 Metallica (2003): „Dirty Window“ (13.1-32.0), L/R-Originalsignal mit L/R-Doppelung der E-Gitarre (oben) sowie M/S-Aufteilung (unten).

Dieses Verfahren lässt sich in vielen Produktionen von Bands mit zwei E-Gitarristen wiederfinden, wo die zwei E-Gitarren, wenn sie von den Spielern links und rechts unisono gespielt werden, klanglich zu einer großen, breiten und fülligen *virtuellen Stereo-E-Gitarre* verschmelzen (vgl. z. B. AC/DC (1980): „You Shook Me All Night Long“, Verse vs. Chorus). Unterschiedliche Gitarren- und Verstärker-Modelle tragen unter Umständen sogar zu noch größeren Differenzen zwischen den beiden Kanälen und somit zu einem noch ausgeprägteren Stereosignal bei. Unabhängig davon, ob die zwei Doubling-Takes von einem Spieler nacheinander oder von zwei Spielern gleichzeitig aufgenommen werden – das Double Tracking Verfahren kann letztlich als eine Spezialform des Overdubbing aufgefasst werden.

HALL- UND ECHO-EFFEKTE

Als ein *stilisiertes* Beispiel dafür, Gesangsaufnahmen bevorzugt in einer reflexionsarmen Gesangskabine anstatt in größeren Studio-Räumen durchzuführen, nennt Simon Zagorski-Thomas den Song „School’s Out“ (1972) von Alice Cooper, produziert von Bob Ezrin. Als wesentlichen Anhaltspunkt für diese Charakterisierung verweist Zagorski-Thomas auf den hochgradig artifiziellen Raumklang, mit dem die Stimme versehen wurde – namentlich ein Slapback-Echo sowie ein kurzer, heller Plate-Hall (vgl. Zagorski-Thomas 2010: 257). Um bezüglich des klanglich tatsächlich sehr auffälligen, schnellen Echos etwas konkretere analytische Aussagen treffen zu können, wird im Folgenden versucht, die Delay-Zeit möglichst exakt zu ermitteln. Dazu ist es zunächst sinnvoll herauszufinden, in welchem Frequenzbereich das Vocal-Delay am deutlichsten in Erscheinung tritt, um es anschließend bestmöglich

isolieren zu können. Daher galt die Suche einer Stelle mit ausgeprägten, hochfrequenten Konsonanten, auf die das Delay präsent folgt. Als geeignet schien ein kurzer Ausschnitt (1:0.2-1:2.5), in dem der Konsonant *t* des Wortes *to* in den Vocals deutlich heraussticht. Abb. 7 enthält zwei separate Darstellungen des Frequenzspektrums unmittelbar vor und dann während dieses Konsonanten. Dem relativ schwach ausgeprägten, stetigen Abfall in den Höhen ab ca. 2 kHz des ersten Beispiels steht ein enormer Impuls im Bereich zwischen 5 kHz und 10 kHz gegenüber, der ausschließlich durch den hier sehr hochfrequenten Konsonanten *t* zustande kommt. Um genau diesen für die Konsonanten relevanten Frequenzbereich für die Analyse zu extrahieren, wurde ein steilflankiger Bandpassfilter zwischen 5 kHz und 10 kHz auf das Originalsignal angewendet und der Pegel der besseren Sichtbarkeit halber angehoben, siehe Abb. 8 (unten). Das auf diese Weise gewonnene Signal liefert einen optimalen Einblick in das Timing und das Feedback des Vocal-Delays in diesem Stück: Der im hochfrequenten Bereich stark ausgeprägte Konsonant *t* schlägt sich in einem enormen Pegelsprung ungefähr in der Mitte der Grafik nieder. Unmittelbar darauf folgt ein etwas kleinerer Wellenberg, danach ein noch kleinerer dritter. Es handelt sich also um zwei in der Lautstärke stetig abnehmende Echos, die – das ergibt eine simple Laufzeitmessung zwischen erstem und zweitem Wellenberg – im Abstand von 85 ms aufeinander folgen. Um einmal mit Notenwerten zu veranschaulichen, wie kurz das Zeitintervall von 85 ms ist: Dies entspricht der Dauer einer Zweiunddreißigstel bei Tempo 88 bpm oder der Dauer einer Sechzehnteltriolen bei Tempo 118 bpm.²

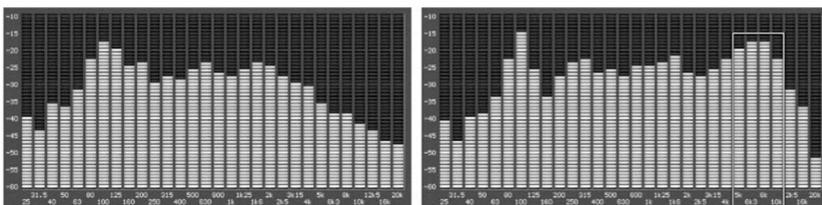


Abb. 7 Alice Cooper(1972): „School’s Out“ (1:0.2-1:2.5), Frequenzspektren unmittelbar vor und während eines hochfrequenten Konsonanten der Vocals.

2 Dies lässt sich anhand folgender Formel berechnen: $\text{Zeitintervall [ms]} = 62.5 * \text{Notenwert [ticks]} / \text{Tempo [bpm]}$, wobei bezüglich des Notenwerts gilt: 960 ticks entsprechen dem Wert einer Viertelnote; 480 ticks einer Achtelnote, 320 ticks einer Achteltriolen usw.

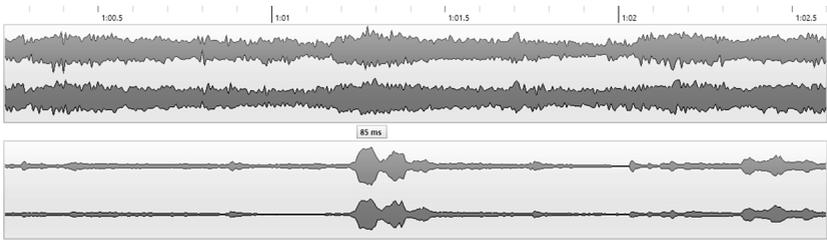


Abb. 8 Alice Cooper (1972): „School’s Out“ – (1:0.2-1:2.5),
oben: L/R-Originalsignal; unten: Bandpassfilterung zwischen 5 kHz und 10 kHz.

Betrachten wir demgegenüber den Anfang der ersten Strophe des Songs „Boulevard of Broken Dreams“ von Green Day (2004) im Ausschnitt 11.5-17.3 Sekunden. Der Song wurde produziert von Rob Cavallo und Green Day, gemischt von Chris Lord-Alge. Um die Wellenformdarstellung zu einem Achtel-Raster zu synchronisieren, wurde einerseits das Tempo mit 83.5 bpm ermittelt und andererseits das Audiosignal entsprechend horizontal ausgerichtet. Um den Fokus auf die Stimme zu richten, erhielten die M- und S-Signale zunächst eine Frequenz-Einschränkung mittels eines steilflankigen Bandpassfilters zwischen 350 Hz und 860 Hz. Eine starke Kompression des S-Signals mit schnellstmöglichen Zeitkonstanten diente darüber hinaus dazu, das Verhältnis zwischen Stimme und den in Sechzehnteln pulsierenden E-Gitarren-Akkorden nochmals zugunsten der Stimme zu optimieren. Das Ergebnis ist in Abb. 9 dargestellt.



Abb. 9 Green Day (2004): „Boulevard of Broken Dreams“ (11.5-17.3),
Visualisierung des Vocal Delays. Oben: L/R Originalsignal; unten: M/S Bandpassfilterung zw. 350 Hz und 860 Hz, das S-Signal wurde zusätzlich noch komprimiert.

Eine eingehende Betrachtung der bearbeiteten M- und S-Signale fördert Bemerkenswertes zutage: Der Ausschnitt beginnt auf Schlag Eins der ersten Strophe; anhand der vertikalen Rasterung lässt sich im Seitensignal wunderbar die für dieses Stück so charakteristisch *zerschnittene* bzw. pulsierende E-Gitarre ablesen. Während der ersten vier Sechzehntel ist im Seitensignal nichts anderes zu sehen; erst ab dem fünften Sechzehntel, also auf Schlag Zwei, erkennt man den einsetzenden Gesang in Form eines durchgezogenen Blocks, der sich unter den Sechzehntelpuls legt. Wenn man das Stück hört – und wie es auch aus dem bearbeiteten Mittensignal hervorgeht –, setzt der Gesang jedoch faktisch bereits auf dem dritten Sechzehntel, also auf Schlag *Eins und ein*. Wie auch im weiteren Verlauf der Grafik ersichtlich wird, hinkt das im bearbeiteten Seitensignal visualisierte Vocal-Delay den eigentlichen Vocals, veranschaulicht im bearbeiteten Mittensignal, stets um den präzisen Wert einer Achtelnote hinterher. Sogar der etwas zu frühe Einsatz spiegelt sich im Delay exakt wider. Außerdem belegt die Visualisierung, dass der Gesang mono und exakt mittig gemischt ist, denn er ist beispielsweise während des dritten und vierten Sechzehntels – also noch bevor das Delay im Seitensignal einsetzt – ausschließlich im Mittensignal und nicht im Seitensignal zu finden. Aus der originalen L/R-Darstellung ist weder die Sechzehntel-Gitarre noch das Vocal-Delay nachvollziehbar; ebenso wenig die Tatsache, dass die Vocals perfekt mittig sitzen. Erst durch eine gezielte Bearbeitung der M- und S-Komponenten offenbaren sich diese an sich verborgenen Details nicht nur auditiv, sondern auch visuell. Man kann schließlich festhalten, dass es sich hier um einen vollkommen anderen Ansatz handelt, Delay zu konfigurieren, als im Song „School’s Out“ mit seinem Mikro-Delay: Ein auf das Tempo des Stücks abgestimmtes und dazu noch im Seitensignal angesiedeltes Delay vermag hier offenbar, zu einer subtilen, im Endeffekt gleichwohl effektiven Unterstützung des Achtel-Grooves beizutragen. Somit enthält der vielschichtige Effekt-Begriff immer auch eine hörpsychologische Dimension.

SCHLUSSBEMERKUNGEN

Mit je nach Stück und Kontext ausgewählten technischen Verfahren, unter anderem namentlich der MS-Aufspaltung, der Hochpass-, Tiefpass- oder Bandpassfilterung und der bedarfsweise visuellen Hervorhebung durch zusätzliche Kompression, konnten auf Basis des bearbeiteten Audio-Signals teils neue oder tiefgreifende Erkenntnisse über Sound gewonnen werden. Durch eine genaue Dokumentation aller Arbeitsschritte lassen sich die daraus gewonnenen Erkenntnisse reproduzieren und nachvollziehen. Anhand diverser Vergleiche früherer und neuerer Stücke konnte ein Wandel in grundlegenden Mixing-Paradigmen nachgewiesen werden, wobei bemerkenswert ist, dass dieser Wandel nicht unbedingt dem technischen Fortschritt zuzuschreiben ist, wie am Beispiel der Mono-Kompatibilität gezeigt wurde. Denn auch viele namhafte Mix Engineers sowie die Ausstattungen moderner, größerer Studios bezeugen, dass in heutigen Mischungen bevorzugt begehrtes Studio-Equipment aus den 1960er und 70er Jahren verwendet wird. Solche *Vintage*-Schätzchen, wie z. B. Urei 1176, Teletronix LA-2A, Fairchild 670, Pultec EQP-1A, Neumann U67 etc., werden auf dem Gebrauchtmarkt teilweise für astronomische Summen gehandelt; einige Hersteller versuchen sich an – klanglich meist unterlegenen – Nachbauten oder Software-Simulationen. Insbesondere Studio-Bilder und Interviews des Mix-Engineers Chris Lord-Alge verdeutlichen, dass er grundsätzlich mit diesen alten, bewährten, teils markant, druckvoll oder auch färbend klingenden Originalen mischt. Somit hätten bereits in den 1970ern praktisch alle Mittel zur Verfügung gestanden, um Mischungen so zu kreieren, dass sie so modern wie heute geklungen hätten. Die enormen Unterschiede lassen sich demnach vielmehr auf einen fundamentalen Wandel in der Klangästhetik zurückführen.

In Bezug auf die gerne bemühte Sound-Box-Methode muss schließlich noch die Frage aufgeworfen werden, ob die Piktogramme überhaupt präzise genug gestaltet werden können, um in ihrer Höhe, Breite, Tiefe, 3D-Positionierung und bildlichen Aussagekraft das tatsächliche Klanggeschehen nachvollziehbar und adäquat zu repräsentieren. Problematisch ist vor allem die Breite, denn diese scheint weder in eindeutiger Korrelation zur Positionierung einer Schallquelle im Stereofeld noch zu deren Stereobasisbreite zu stehen. Eine Mono-Schallquelle kann strenggenommen, außer durch einen Punkt, nicht adäquat durch Bilder mit einer gewissen Breite ohne eindeutig definierten Mittelpunkt abgebildet werden; umgekehrt suggeriert jede grafisch unumgängliche Breite eine entsprechende Ausprägung in der Stereointensität, was bei Mono-Signalen jedoch irreführend ist. Eine dezidierte Kritik der Sound-

Box-Methode muss jedoch einem weiteren, bereits entstehenden Forschungsbeitrag vorbehalten bleiben, da eine umfangreiche Würdigung diesen Rahmen sprengen würde. Ein Anliegen dieses Beitrags war es jedenfalls zu zeigen, dass gerade im stereofonen Bereich unter Umständen andere Methoden und Herangehensweisen, teils andere Paradigmen geeigneter sind, um ein moderneres Klanggeschehen zu erfassen. An diese Erkenntnisse anknüpfend ließen sich noch Untersuchungen anstellen, die weitere Aspekte der produktionstechnischen Klanggestaltung in den Blick nähmen, wie z. B. Dynamik-Kompression, Sidechain- oder Summen-Kompression, Mastering, Lautheit etc. Außerdem könnte der Frage nachgegangen werden, inwieweit bestimmte klangliche Charakteristika auf einen spezifischen Personalstil zurückzuführen sind (z. B. *der Sound des Chris Lord-Alge*), oder eher ein zeitliches und/oder Genre-spezifisches Phänomen darstellen. Ein diesbezügliches Forschungsprojekt befindet sich bereits im Entstehungsprozess.

BIBLIOGRAFIE

- Bartholomew, Wilmer T.: A Physical Definition of ›Good Voice-Quality‹ in the Male Voice. In: *The Journal of the Acoustical Society of America* 6 (1) 1934, S. 25-33.
- Bartlett, Bruce/ Bartlett, Jenny: *Practical Recording Techniques*. Dritte Edition. London: Focal Press 2002.
- Berger, Harris M./ Fales, Cornelia: ›Heaviness‹ in the Perception of Heavy Metal Guitar Timbres: The Match of Perceptual and Acoustic Features over Time. In: Paul D. Greene/ Thomas Porcello (Hg.): *Wired for Sound: Engineering and Technologies in Sonic Cultures*. Middletown: Wesleyan University Press 2005, S. 181-197.
- Braae, Nick: The Development Of The ‘Epic’ Queen Sound. In: *Journal on the Art of Record Production* (10) 2015.
- Bunge, Ernst: *Vergleichende systematische Untersuchungen zur automatischen Identifikation und Verifikation kooperativer Sprecher*. Diss. Technische Hochschule Darmstadt 1977.
- Camilleri, Lelio: Shaping sounds, shaping spaces. In: *Popular Music* 29 (2) 2010, S. 199-211.
- Dickreiter, Michael u. a. (Hg.): *Handbuch der Tonstudioteknik*, Bd. 1 & 2. Walter de Gruyter GmbH 2014.
- Dockwray, Ruth/ Moore, Allan F.: Configuring the sound-box 1965-1972. In: *Popular Music* 29 (2) 2010, S. 181-197.

- Eddins, David A./ Green, David M.: Temporal Integration and Temporal Resolution. In: Brian C. J. Moore (Hg.): *Hearing. Handbook of Perception and Cognition*. Zweite Ausgabe. San Diego: Academic Press 1995, S. 207-242.
- Födermayr, Franz: Populärmusik als Gegenstand musikwissenschaftlicher Forschung. Ein hermeneutischer Versuch. In: *Musicologica Austriaca* 5 1985, S. 49-84.
- Födermayr, Franz/ Deutsch, Werner A.: Zum stimmlichen Ausdrucksrepertoire von Hank Williams. In: Helmut Rösing (Hg.): *Rock, Pop, Jazz – musikimmanent durchleuchtet*. Hamburg: CODA-Verlag 1989, S. 93-105.
- Großmann, Rolf: Signal, Material, Sampling. Zur ästhetischen Aneignung medientechnischer Übertragung. In: Sabine Sanie/ Christian Scheib (Hg.): *Übertragung – Transfer – Metapher. Kulturtechniken, ihre Visionen und Obsessionen*. Bielefeld: Kerber 2004, S. 91-110.
- Hähnel, Tilo/ Marx, Tobias/ Pfleiderer, Martin: Methoden zur Analyse der vokalen Gestaltung populärer Musik. In: *Samples* 12 2014.
- Holland, Michael: Rock Production And Staging In Non-Studio Spaces: Presentations Of Space In Left Or Right’s Buzzy. In: *Journal on the Art of Record Production* (8) 2013.
- Lacasse, Serge: Slave to the Supradiegetic Rhythm: A Microrhythmic Analysis of Creaky Voice in Sia’s ›Breathe Me‹. In: Anne Danielsen (Hg.): *Musical Rhythm in the Age of Digital Reproduction*. Ashgate Publishing 2010, S. 141-156.
- Michelsen, Morten: Michael Jackson’s sound stages. In: *Sound Effects* 2 (1) 2012, S. 82-95.
- Moore, Allan F.: The Textures of Rock. In: Rosanna Dalmonte/ Mario Baroni (Hg.): *Secondo Convegno Europeo di Analisi Musicale*. Trient 1992, S. 241-244.
- Moore, Allan F.: *Rock: the primary text. Developing a musicology of rock*. Aldershot: Ashgate 2001.
- Moore, Allan F.: Introduction. In: Allan F. Moore (Hg.): *Analyzing Popular Music*. Cambridge University Press 2003, S. 1-15.
- Moylan, William: *The Art of Recording: The Creative Resources of Music Production and Audio*. New York: Van Nostrand Reinhold 1992.
- Moylan, William: Considering Space In Music. In: *Journal on the Art of Record Production* (4) 2009. Online unter: <http://arjournal.com/1379/considering-space-in-music/> (Zugriff am 25. 10. 2014).

- Robjohns, Hugh: How do I create a stereo mix from mono material? In: *Sound On Sound* 2003 (12). Online unter: <http://www.soundonsound.com/sos/dec03/articles/qanda.htm> (Zugriff am: 06. 03. 2015).
- Schneider, Albrecht: Klanganalyse als Methodik der Populärmusikforschung. In: Helmut Rösing/ Albrecht Schneider/ Martin Pfeleiderer (Hg.): *Musikwissenschaft und populäre Musik: Versuch einer Bestandsaufnahme* (Hamburger Jahrbuch für Musikwissenschaft Bd. 19). Frankfurt am Main u. a.: Lang 2002, S. 107-129.
- Sundberg, Johan: Formant frequencies of bass singers. In: *STL-QPSR* 9 (1) 1968, S. 1-6.
- Voß, Ingo: Pseudostereofonie. In: *Funkschau Heft* 6 1977, S. 245-248.
- White, Paul: Improving your stereo mixing. In: *Sound On Sound* 10/2000. Online unter: <http://www.soundonsound.com/sos/oct00/articles/stereo mix.htm> (Zugriff am: 09. 03. 2015).
- Winckel, F.: Physikalische Kriterien für objektive Stimmbeurteilung. In: *Folia Phoniatica et Logopaedica* 5 (4) 1953, S. 232-252.
- Zagorski-Thomas, Simon: The stadium in your bedroom: functional staging, authenticity and the audience-led aesthetic in record production. In: *Popular Music* 29 (2) 2010, S. 251-266.
- Zeiner-Henriksen, Hans T.: Moved by the Groove: Bass Drum Sounds and BodyMovements in Electronic DanceMusic. In: Anne Danielsen (Hg.): *Musical Rhythm in the Age of Digital Reproduction*. Ashgate Publishing 2010, S. 121-140.

DISKOGRAPHIE

- AC/DC: You Shook Me All Night Long, auf: *Back in Black*. LP. Atlantic 1980.
- Beach Boys: *Pet Sounds*. LP. Capitol Records 1966.
- Checkmates Ltd., The: *Love Is All We Have To Give*. LP. A&M Records 1969.
- Cooper, Alice: School's Out, auf: *School's Out*. LP. Warner Bros, 1972.
- Daughtry: Over You., auf: *Daughtry* (Album). CD: RCA Records, 2006.
- Green Day: American Idiot, Boulevard of Broken Dreams auf: *American Idiot*. CD. Reprise Records, 2004.
- Humperdinck, Engelbert: There Goes My Everything [1967], auf: *Engelbert: His Greatest Love Songs*. CD. Universal Music 2004.

Kinks, The: Lola, auf: *Lola Versus Powerman and the Moneygoround, Part One*. LP. Pye/Reprise, 1970.

Metallica: Dirty Window, auf: *St. Anger*. CD: Elektra Records, 2003.