

# **Neuronale Koinzidenz mit Unschärfe löst das Problem unbefriedigender Konsonanztheorien**

Jobst Peter Fricke

## **Zusammenfassung**

Es wird vorgeschlagen, Lickliders Autokorrelationstheorie der Tonhöhenwahrnehmung (1951, 1959), die Hesse in Form der Zeitreihen-Korrelationsanalyse auch zur Erklärung der Konsonanz heranzog (1972, S. 154; 2003, S. 142), mit einem Zeitfenster für die Unschärfe der Gleichzeitigkeit zu versehen, um sie dadurch praxiskompatibel zu gestalten. Bisherige Konsonanztheorien basieren auf der Darbietung stationärer Klänge und können die Verstimmungstoleranz nicht erklären. Die in der frequency domain operierenden Konsonanztheorien sind noch weniger dazu in der Lage, die Bedingungen der praktischen Musikausübung zu berücksichtigen, als die in der time domain prozessierenden. Mit Langners Nachweis einer doppelten Repräsentation von Tonhöhe, der tonotopen und der periodotopen Abbildung sowohl im Zentrum des Colliculus Inferior (ICC) als auch im Cortex (Langner et al., 1988, 1997), sowie mit den Interspike-Histogrammen von Tramo et al. (2001), die zeigen, dass mit neuronaler Autokorrelation harmonische Intervalle wegen ihrer Periodizität erkannt werden, sind die Voraussetzungen dafür geschaffen, Konsonanztheorien in der time domain zu favorisieren. Denn es ist erstmals die Existenz solcher neuronalen Prozesse nachgewiesen, die dazu notwendig sind, Konsonanz mit der Periodizitätsanalyse der neuronalen Impulsmuster zu erklären.

Die Rigidität der Autokorrelationsanalyse, die zur Folge hat, dass verstimmte konsonante Intervalle nicht als konsonant erkannt werden, obwohl sie (noch) konsonant klingen, wird beseitigt durch die Einführung einer Unschärfe beim Erkennen der Gleichzeitigkeit. Sie ist mit der endlichen Breite der Nervenimpulse sowie der endlichen Zeit der Reaktion auf das fast gleichzeitige Eintreffen zweier Nervenimpulse an einem Neuron zu begründen (Fricke, 2005a, S. 135). Mit der Berücksichtigung dieser Unschärfe durch ein Zeitfenster von 0,8 ms, mit der Ebeling die Autokorrelation der Intervalle kürzlich berechnet hat, ergibt sich eine Abstufung der konsonanten Intervalle in der Rangfolge, wie sie Stumpf (1890, S. 176) aufgrund seiner Verschmelzungsuntersuchungen angegeben hat. Die Kurve zeigt insbesondere, dass jedes Intervall sich mit der notwendigen Intervallbreite darstellt,

die durch die Verstimmungstoleranz gegeben ist. Damit ist es erstmalig gelungen, das konsonante Zusammenklingen auch verstimmter Intervalle innerhalb einer Konsonanztheorie zu erklären.

## Abstract

It is suggested to provide Licklider's autocorrelation theory of pitch perception (1951, 1959), that Hesse also used as time series analysis of correlation for an explanation of consonance, with a window for the uncertainty of simultaneousness, thus making it a practicable method. Current consonance theories are based on the presentation of steady sounds and cannot explain the tolerance of detuning. The consonance theories operating in the frequency domain are even less able to consider the terms of music practice than the time domain based theories. Regarding Langner's proof of the dual representation of pitch, the tonotopic and periodotopic mapping in the center of the colliculus inferior (ICC) as well as in the cortex (Langner et al., 1988, 1997), and the interspike-histograms of Tramo et al. (2001), which demonstrate that, using neural autocorrelation, harmonic intervals are detected because of their periodicity, a supposition to favour theories in the time domain is established. For the first time the existence of this kind of neural processes could be proved, that are necessary to explain consonance by means of a periodicity analysis of the neural pulse pattern.

The rigidity of the autocorrelation analysis, that results in identifying detuned consonant intervals as not consonant, even if they (still) sound consonant, is eliminated by the introduction of an uncertainty connected with the detection of simultaneousness. This can be explained by the finite width of the nerve impulses and the finite reaction/latencies concerning the almost simultaneous arrival of two nerve impulses at one neuron (Fricke, 2005a, S. 135). Considering that uncertainty through a 0.8 ms window, like Ebeling currently did for calculating the autocorrelation of intervals, a graduation of the consonant intervals according to the ranking order of Stumpf's (1890, S. 176) investigations of tonal fusion is obtained. The curve particularly shows that each interval has the necessary interval width, that is given by the tolerance of detuning. It therewith succeeded for the first time to explain the consonant simultaneous-sounding of even detuned intervals within a theory of consonance.

## 1 Einleitung

Unter den in der Musik verwendeten Intervallen gelten seit mindestens 2.500 Jahren diejenigen als konsonant, die durch einfache Längenverhältnisse von schwingenden Saiten definiert sind: Oktave  $2/1$ , Quinte  $3/2$ , Quarte  $4/3$ . Zur Begründung dieses Zusammenhangs wurden seither zahlreiche Theorien

aufgestellt, die hinsichtlich Beweiskraft und Gültigkeitsanspruch heutigen Vorstellungen nicht mehr genügen. Sie lassen sich in drei Gruppen einteilen, von denen die eine die spekulativen Theorien umfasst, während die anderen beiden erklärende Theorien enthalten.

1. Die Theorie der *pythagoräischen Schule*, die später ihrem Protagonisten, dem Pythagoras selbst zugeschrieben wurde, ist als eine solche, die die Phänomene nicht erklärt, sondern nur beschreibt, eine so genannte spekulative Theorie. Aber sie teilt erstmals ein Gesetz mit, das ein von Musikern beobachtetes Phänomen in Regeln fasst und mit physikalischen Mitteln darstellt (Békésy, 1962). Die Mittel sind die Längen von Saiten. Die Verhältnisse dieser Saitenlängen sind den verschiedenen Intervallen zugeordnet. Die Gesetzmäßigkeit besteht darin, dass diese Verhältnisse unabhängig sind von der Spannung der Saiten, dem Material, dem Querschnitt und den gewählten Längen, die die Proportionen bilden.
2. Theorien, die in der *time domain* Erklärungen für die Besonderheiten des Zusammenklings konsonanter Intervalle suchen, und
3. Theorien, die in der *frequency domain* konsonantes Verhalten zu erklären versuchen.

Bei den erklärenden Theorien (Fälle 2 und 3) wurde zur Begründung zunächst auf die geordnete Bewegung beim Zusammentreffen der Schwingungen der beiden Töne im Ohr verwiesen (Beeckmann, 1614/15 nach Muzzolini, 2006, S.158; Galilei, 1638; Euler, 1739). Sie blieben aber die Erklärung schuldig, wie diese Ordnung erkannt werden könnte und warum geordnete Bewegung für das Hören der Konsonanz von Vorteil sei. Später wurde das Zusammentreffen (die „Koinzidenz“) möglichst vieler Obertöne der beiden Intervalltöne für Konsonanz verantwortlich gemacht (Rameau, 1722). Diese Obertöne verursachen im Falle der Dissonanzen, aber auch bei verstimmten Konsonanzen, Schwebungen, die, wenn sie schnell aufeinander folgen, zu Rauigkeit führen. Schwebungen und Rauigkeit als Unterscheidungsmerkmal für Kon- und Dissonanz zu nutzen, ist das Kennzeichen der sogenannten Störungstheorien (Helmholtz, 1863; Plomp & Levelt, 1965). Sie waren gegenüber den anderen Theorien über hundert Jahre im Vorteil, da sich die zu ihrer Begründung notwendigen Phänomene anatomisch und physiologisch nachweisen ließen. Es gibt allerdings schwerwiegende Einwände, die folgendermaßen zu begründen sind:

1. geringfügige Verstimmungen von konsonanten Intervallen lassen diese nicht dissonant werden, obwohl die berechneten Störungen erheblich sind. Die Grenze maximaler Rauigkeit würde bei obertonreichen Klängen mit 12 bis 18 Teiltönen bei einer Quinte (c,g) theoretisch schon bei 0,5 % entsprechend 8 Cents Verstimmung erreicht, weil zwischen den Obertönen ein Vielfaches der Differenz auftritt, die zwischen den Grundtönen besteht. (Man beachte: als Richtwert für den ebenmerklichen Unterschied gilt 0,3 % entsprechend 5 Cents.) Im Gegensatz zu dieser angeblichen, nur

auf Berechnung beruhenden Störung gibt es experimentell nachgewiesen eine beachtliche Toleranz (bis 25 Cents =  $\frac{1}{4}$  Halbton temperierter Stimmung) je nach instrumentaler Ausführung und Position in der musikalischen Komposition (Shackford, 1962, S. 300; Fricke, 1973, S. 398; Fricke, 2005b, S. 591).

2. Die instationären Vorgänge der „normalen“ Klangerzeugung schwächen bzw. unterdrücken das Phänomen Schwebung und Rauigkeit. Je nach Art und Zweck der Klangerzeugung ist davon auszugehen, dass es erwünschte und unerwünschte Formen instationärer Vorgänge bei der Klangerzeugung gibt. Gerade bei perfekter, optimaler Klangerzeugung ist in der musikalischen Praxis in jedem Falle von instationärer Tongebung auszugehen; denn die Klänge werden in diesem Fall in der guten Absicht, sie „natürlicher“ und „lebendiger“ erklingen zu lassen, mit natürlichen Fluktuationen und darüber hinaus sogar mit Vibrato erzeugt. Das hat zur Folge, dass – insbesondere bei mit Vibrato erzeugten Tönen – Schwebungen nicht entstehen oder entstehende Schwebungen nicht hörbar sind. Unter physikalisch/physiologischem Gesichtspunkt betrachtet sind Störungen (eines gleichmäßigen Ablaufs der Tonerzeugung) in der Musik allgegenwärtig und Störungen treten bei der Verstimmung eines konsonanten Intervallverhältnisses in den Hintergrund bzw. gar nicht erst auf. Bei synthetischer und dadurch oft allzu starrer Klangerzeugung fanden langsame Schwebungen aufgrund von Verstimmungen sogar positive Beurteilungen, da auf diese Weise der Klang durch die Fluktuationen angenehmer empfunden wird (Miskiewicz et al., 2003). Damit aber kommt ein Gesichtspunkt in die Beurteilung, der mit dem Konsonanzempfinden nur nebensächlich zu tun hat: der der angenehmen Empfindung.

Die Skala angenehm/unangenehm läuft in vielen Fällen mit der Beurteilung konsonant/dissonant parallel. Sie wurde sogar zu einem Gradmesser für Konsonanz eingesetzt (Plomp & Levelt, 1965). Wie verfehlt diese Sichtweise aber ist, zeigen gerade die Ergebnisse dieser Studie von Miskiewicz und Rogala (2003). Denn es ergab sich das paradoxe Ergebnis: Verstimmung aus den ganzzahligen Intervallverhältnissen ist bei stationären Klängen zwecks Fluktuationsanreicherung eine erwünschte Maßnahme.

Keine der oben unter 1 bis 3 aufgeführten Theorien ist also in der Lage, ihre Gültigkeit und ihren Geltungsbereich auch unter den Bedingungen der praktischen Musikausübung zu erklären. Konsonanz darf einerseits nicht durch den physikalischen Sachverhalt und dessen mathematische Beschreibung in Zahlenverhältnissen, andererseits aber auch nicht durch die physiologischen bzw. neuropsychologischen Prozesse auf der Basilarmembran und in den Nervenschaltvorgängen definiert werden, sondern umgekehrt: die durch den Gebrauch in der Musik hervorgehobene Stellung konsonanter Intervalle kann möglicherweise und in gewissem Umfang mit den physikalischen und physiologischen Vorgängen begründet werden. Ein Versuch in dieser Richtung, der sich in weitgehender Übereinstimmung mit dem derzeitigen Stand der Kenntnisse befindet, wird hiermit unternommen.

Die auf subharmonische virtuelle Spektralkomponenten gegründete Theorie Terhardts (1976) findet in der neurophysiologischen Auswertung der Schallereignisse bislang keine Entsprechung. Auch sind die von dieser Theorie geforderten Prinzipien ganzheitlich-gestalthafter Wahrnehmung und gestaltpsychologischer Prozesse, wie sie beim Hören von Akkorden zu beobachten sind (Louven, 2005, S. 209), auf so elementarer Ebene wie dem Zusammenspiel von subharmonischen Spektralkomponenten nicht bekannt geworden. Es gibt allerdings Fälle der Entstehung von Residualtönen, die auf dem von Terhardt (1972) vorgeschlagenen Verfahren beruhen könnten. Sie entstehen z. B., wenn man bei einer Pfeifenorgel die Aliquote Quinte  $2\frac{2}{3}'$ , Oktave  $2'$  und Terz  $1\frac{3}{5}'$  in einer Melodie parallel laufen lässt, ohne dass der Grundton  $8'$  und Oktave  $4'$  erklingt. Es stellt sich dann erst nach einigen Tönen die Grundtonempfindung ein. Bei der „normalen“ Residualtonbildung (Schouten, 1940a, 1940b; Meyer-Eppler et al., 1959; Walliser, 1969), die wir aus dem täglichen Umgang kennen (Telefon etc.), tritt die Grundtonempfindung immer spontan und unmittelbar auf. Es liegt bei den Aliquot-Tönen, wenn sie in verschiedene Frequenzgruppen fallen und nur deren Teiltöne sich in den Frequenzgruppen zur gemeinsamen Periodenbildung überlagern, offenbar ein anderer kognitiver Prozess vor. Hier gibt es also noch Forschungsarbeit zu leisten.

Mit der von Licklider (1951, 1959) vorgeschlagenen Autokorrelationstheorie für die Tonhöhenwahrnehmung wurde das Augenmerk wieder mehr auf die neuronale Verarbeitung in der time domain gelenkt, die nach (den Mikrorhythmentheorien) von Lipps (1899) und von Hornbostel (1926) in den Hintergrund getreten war.

Seit Langner (ab 1988) mit der periodotopen Abbildung im ICC und in der Hirnrinde nachgewiesen hat, dass die anatomischen und neurophysiologischen Voraussetzungen für die Theorien der time domain vorhanden sind, und seit Tramo et al. (2001) die Autokorrelationstheorie Lickliders in neurophysiologischen Prozessen nachgewiesen haben, fehlte für eine schlüssige Theorie der Konsonanz nur noch, dass diese Autokorrelationstheorie auch für verstimmte konsonante Intervalle die konsonante Wirkung noch realitätsgerecht und praxiskonform einstuft und damit die Verstimmungstoleranz des Hörens berücksichtigt. Es fehlte bis jetzt eine Konsonanztheorie, die auch praxiskompatibel ist.

## 2 Neue Argumente für eine alte Theorie

Beide derzeit favorisierte Theorien sind unter Berücksichtigung der Erfahrungen bei der praktischen Musikausübung also nicht überzeugend. Die auf Schwebungen gegründete Theorie kann Konsonanz nicht erklären, weil Schwebungen in der praktischen Musikausübung keine das Zusammenklingen regulierende und die Konsonanz von der Dissonanz unterscheidende Rolle spielen. Ihre Hilfestellung beim Einstimmen von Musikinstrumenten (Klavier, Orgel, Streich- und Blasinstrumente) ist unbestritten. Diese Hilfs-

funktion, die nur unter besonderen Bedingungen zur Geltung kommt, wurde fälschlicherweise zur Hauptfunktion und als solche für die Unterscheidung von Konsonanz und Dissonanz herangezogen. Dafür gibt es historische Gründe, die in der damals noch nicht so weit entwickelten apparativen Ausrüstung zu suchen sind. So galt z. B. das Harmonium, dessen Töne wegen ihres stationären Verhaltens und der zahlreichen Obertöne die Schwebungen besonders deutlich hervortreten lassen, wenn Intervalle verstimmt oder unharmonisch sind, als ideales Demonstrationsinstrument für Tonsysteme und die darin auftretenden Intervallproportionen (Fricke, 2002, S. 105–116). Neurophysiologische Befunde im Zusammenhang mit dem Stand der Erkenntnis über die Natur musikalischer Klänge zeigen hingegen, dass die Schwebung die ihr zugeschriebene Bedeutung in diesem Zusammenhang nicht haben kann. (Schwebungen sind eine in der praktischen Musikausübung zu selten zu beobachtende Erscheinung. Sie können für ein so herausragendes Phänomen wie das konsonierende Zusammenklingen von Tönen keine ausreichende Begründung geben).

Es gibt inzwischen aber genügend neurophysiologische Belege für jene Theorie, die zu den ältesten der erklärenden Theorien gehört: die Zeitmustertheorie, die auf „Zeitreihen-Korrelationsanalyse“ beruht (Hesse, 2003, S. 142). Sie kann das Phänomen des konsonanten Zusammenklagens auch unter den Bedingungen der praktischen Musikausübung erklären, wenn eine Unschärfe der neuronalen Verarbeitung berücksichtigt wird. Sie beruht auf einer Zeitspanne, die für die neuronale Reaktion der Nervenzelle und für die endliche zeitliche Breite der Nervenimpulse (spikes) in der Autokorrelationsanalyse anzusetzen ist (Fricke, 2005a, S. 134–135). Diese Unschärfe resultiert einerseits schon aus der logischen Konsequenz, dass neuronale Schaltprozesse nicht mit elektronischen Simulationen gleichzusetzen sind, sondern reale Vorgänge des Ionenaustauschs in Nervenzellen sind. Deshalb lieferten Modellrechnungen mit Zeitreihenanalysen auf der Basis der Autokorrelation bisher unbefriedigende Ergebnisse, die die Erfahrungen mit der praktischen Musikausübung nicht abbilden konnten. Hinzu kommen in den letzten Jahren Belege für die Existenz und die Größe der Unschärfe aus der neurologischen Forschung (Dudel, 1990, S. 52; Tramo et al., 2001).

## *2.1 Koinzidenz in der frequency domain*

Störungstheorien erklären Konsonanz mit der möglichst weitgehenden Abwesenheit von Schwebungen zwischen Tönen, deren Frequenzen in einem konsonanten Intervallverhältnis stehen. Sie operieren im Spektralbereich, d. h. in der frequency domain, und betrachten dort die harmonischen Obertöne der beiden Intervalltöne. Es kommt dabei auf die Positionen der Obertöne an, insbesondere aber, wie viele der über den verschiedenen Grundtönen stehenden Obertöne auf der Frequenzachse die gleiche Position einnehmen und welche Abstände sich zwischen ihnen ergeben. Grundsätzlich wird dabei zunächst von der harmonischen Struktur der Obertonverteilung und damit

von der periodischen zeitlichen Struktur der Schwingungen stationärer Töne ausgegangen. Die gelegentlich in dieses Konzept einbezogenen Kombinationstöne fügen sich dem harmonischen Gerüst der Obertöne vollständig ein, wenn die Primärtöne (das Ausgangsintervall) in einem harmonischen Verhältnis stehen. Die Kombinationstöne werden dann hilfsweise zur Unterstützung des Konzepts der für die Theorie erforderlichen harmonischen Struktur herangezogen, wenn die notwendige Zahl der Obertöne fehlt. Es gibt inzwischen aber genügend experimentelle Hinweise dafür, dass ihre Bedeutung überschätzt wurde (Fricke, 1996).

Im Falle harmonisch reiner Intervalle haben folgende (harmonischen) Obertöne gemeinsame Positionen, sie koinzidieren. Bei der Oktave (C,c) fällt jeder zweite Teilton des tieferen Tones (c, c', g' ...) mit jedem Teilton des höheren Tones (c, c', g', c'', e'' ...) zusammen. Bei der Quinte ist es jeder dritte Teilton des tieferen, der mit jedem zweiten Teilton des höheren zusammentrifft. Bei jedem Intervall  $m:n$  mit  $f_2:f_1 = m:n$  sind es die Frequenzen  $mf_1 = nf_2$  und deren Vielfache ( $m, n$  ganzzahlig), bei denen Koinzidenz eintritt. Setzt man diese Schwingungszahlen  $mf_1, nf_2$  ( $f_2 > f_1, m, n = 1, 2, 3, \dots$ ) für die Töne ein, dann stellt man fest, dass sich beide Reihen in eine gemeinsame harmonische Reihe einfügen, die auf dem virtuellen Grundton aufbaut und rechnerisch durch den größten gemeinsamen Teiler (g. g. T.) bestimmt ist.

Als der harmonische Obertonaufbau der Töne durch Mersenne (1636) bekannt wurde, reizte diese offenbar aufregende Entdeckung nachfolgende Generationen von Musiktheoretikern, darauf die Konsonanztheorie zu gründen. Rameau (1722) entwarf mit Hilfe von d'Alembert auf der Grundlage dieser Erkenntnis, die im Orgelbau jedoch schon lange praktisch umgesetzt worden war, eine Lehre der musikalischen Harmonie, des *basse fondamentale*. Allerdings blieb diese Theorie in der Erklärung der Wirksamkeit dieser Zusammenhänge für das Hören, d. h. hier für das konsonante Zusammenklingen von Tönen im Spekulativen verhaftet.

In den höheren Regionen des Obertonaufbaus kommen die Obertöne der beiden Intervalltöne sich immer näher, die Intervalle werden enger, und zwar um so früher, je weniger konsonant die Intervalltöne sind. Für die Rangfolge der Konsonanz galt schon seit der pythagoräischen Zeit, dass die Oktave eine Ausnahmestellung an der Spitze der Rangfolge einnimmt, dann die Quinte und die Quarte folgen. In den Kreis der Konsonanz wurde von Franco von Köln (1260) dann auch der Didymische Ditonus  $5/4$  als große Terz sowie die kleine Terz  $6/5$  aufgenommen. Spektralkomponenten, die bis auf diese Tonabstände zusammenrücken, bilden – wenn die Intervalltöne nur genügend regelmäßige Schwingungen ausführen – Schwebungen. Das gilt für die dann folgenden kleineren Intervalle, Ganztöne  $9/8, 10/9$ , sowie die Halbtöne  $16/15, 25/24$  usw. und deren Ergänzungen zur Oktave  $16/9, 9/5, 15/8, 12/25$  usw. in verstärktem Maße, da dann zwei und sogar noch mehr Obertöne in eine Frequenzgruppe (critical bandwidth) fallen.

Helmholtz (1863) erkannte diese Zusammenhänge und hatte die Idee, sie zu einer Konsonanztheorie zu entwickeln. In ihr nimmt die Störung durch

Schwebungen eine zentrale Stellung ein. Störend sind nach seinen Angaben besonders die Schwebungen mit der Häufigkeit von 30 bis 40 Pulsen pro Sekunde, wobei im unten angrenzenden Bereich die Störung eher als modulatorische Schwankung, im oberen mehr als Rauigkeit hörbar ist.

Geprüft wurden Schwebungen unter verschiedenen Versuchsbedingungen, zunächst mit Sinustönen in allen Tonhöhenregionen, einerseits auf ihre Hörbarkeit, andererseits auf ihre Lästigkeit. Für Töne mit harmonischen Obertönen kann von dieser Basis aus ein Störungsgrad errechnet werden, wie Plomp und Levelt (1965) es getan haben. Man muss dazu von einer sinnvoll geschätzten, aber doch willkürlichen Annahme über die Zahl und die Stärke der Obertöne ausgehen, so dass das Ergebnis als sehr hypothetisch anzusehen ist. Es bestätigt aber qualitativ die seit der Antike gültige Rangfolge der konsonanten Intervalle. Folgendes sei dazu noch einmal hervorgehoben: Sowohl die Sinustöne selbst als auch die aus ihnen zusammengesetzten harmonischen Spektren stellen streng stationäre und damit periodische Schallvorgänge dar. Wie an Harmoniumtönen, die sich durch einen starren Klang mit einer großen Zahl von Harmonischen auszeichnen, gezeigt werden kann, sind die Störungen erheblich, wenn die Obertöne bei nicht-konsonanten Intervallen zu häufig und zu nah aneinander rücken. Andererseits sind die stationären Klänge, vor allem die synthetisch erzeugten, so starr und unbefriedigend (um nicht zu sagen langweilig), dass Schwebungen als eine willkommene Bereicherung angesehen werden, wenn konsonante Intervalle ein wenig verstimmt werden. Es hängt also sehr von den Ausgangsbedingungen ab, ob und wie Schwebungen zu werten sind. Miskiewicz und Rogala (2003) kamen aus diesen Gründen zu einer positiven Bewertung der Schwebung bei leicht verstimmtten Intervallen (siehe dazu die oben (S. 97) vorgenommene Kommentierung dieser Versuche).

Die in der praktischen Musikausübung verwendeten Klänge besitzen, von wenigen Ausnahmen abgesehen, diese Starrheit nicht. Sie verfügen über eine natürliche Lebendigkeit, die auf Mikromodulationen zurückzuführen sind. Diese im Oszillogramm und Spektrogramm sichtbaren Fluktuationen, die als minimale Frequenz- und Phasenschwankungen der Obertöne sichtbar sind und insgesamt eine Klangfarbenmodulation bewirken, hängen mit den nicht vollständig stationären Bedingungen der Klangerzeugung zusammen. Die Widerhaken am Streichbogen sind statistisch verteilt und die Wirbelbildung an labialen Blasinstrumenten gehorcht stochastischen Gesetzen. Der Musiker, der diese naturgegebene Lebendigkeit des Klangs noch bereichern will, erhöht durch Vibrato die Unregelmäßigkeit der Klangerzeugung. Ein Maximum an statistischen Schwankungen zeigen Töne mit chorischem Effekt. Das Zusammenwirken vieler Töne mit jeweils eigener Schwankungsbreite führt dazu, dass Schwebungen bei verstimmtten Konsonanzen nicht mehr hörbar sind. Solche Töne bestehen selbst aus statistisch schwankenden Schwingungszügen, die aus den stochastischen Anregungsbedingungen und möglicherweise noch zusätzlichen Vibratoschwankungen zusammengesetzt sind; sie stellen selbst statistisch schwankende Schwebungen dar.



Es ergibt sich damit folgendes Bild:

Schwebungen und Rauigkeit sollen (weil sie als störend empfunden werden bei nicht-konsonanten Verhältnissen) verantwortlich sein für die herausgehobene Position konsonanter Intervalle. (Negativ-Definition der Konsonanz).

Die verstärkt auftretenden Phänomene Schwebungen und Rauigkeit bei wenig verstimmten Konsonanzen passen aber nicht in das Konzept der Störungstheorien. Denn die „Störungen“ beeinträchtigen den Eindruck konsonierenden Zusammenklagens nur wenig oder gar nicht.

Andererseits werden Vibrato und statistische Schwankungen als Bereicherung der Tonerzeugung empfunden. Bei starren synthetischen Klängen werden Schwebungen daher als wohltuend eingestuft (Miskiewicz & Rogala, 2003). Daher sollte dieses nicht-schlüssige Konzept der Störungstheorien endgültig aufgegeben werden. Die Negativ-Definition der Konsonanz wird inzwischen auch von Tramo et al. (2001, S. 96) abgelehnt: „consonance cannot be explained solely by the absence of roughness.“

Schwebungen und Rauigkeit können in besonderen Fällen sehr wohl auch in der praktischen Musikausübung eine ganz offensichtliche Wirkung ausüben und erhebliche Bedeutung gewinnen. Dazu zwei Beispiele:

1. Bei Dissonanzen in tiefer Tonlage: Je nach Oberton Aufbau der Klänge können Rauigkeiten, besonders in tiefen Tonlagen, den Dissonanz-Eindruck verstärkend oder auch mildernd wirken. Wolfgang Voigt (1985) hat dies an musikalischen Beispielen nachgewiesen und gezeigt, dass Komponisten diesem Effekt in ihren Kompositionen in Satzweise und Instrumentation durchaus Rechnung tragen.
2. Bei Orgelstimmungen: Durch die Umstimmung einer Orgel wurde das Klangbild hinsichtlich der Rauigkeit verändert. Die Orgel im Freiburger Dom wurde zum Vergleich einmal „wohltemperiert“ (annähernd gleichstufig), einmal in einer der Silbermann-Stimmung ähnlichen Temperatur mit gleichem Repertoire gespielt; letztere Stimmung enthält in bevorzugten Tonarten annähernd reine Terzen (Wegscheider, 1986). In den melodischen Linien waren Unterschiede in den Intervallgrößen (der sukzessiven Intervalle) nicht merkbar, aber an den akkordischen Stellen, und dort ausschließlich bei vollem Werk, war der Klang der eher gleichstufig temperierten Stimmung rauher, der in der Stimmung nach Silbermann weniger rau.

Diese Wirkung von Schwebungen und Rauigkeit in der praktischen Musikausübung bleibt daher unbestritten; sie kann jedoch nicht zur Begründung der Konsonanz herangezogen werden.

## *2.2 Das Geschehen auf der Basilarmembran*

Die Abläufe der Bewegungen auf der Basilarmembran zu kennen, ist entscheidend für das Verständnis aller weiteren Stufen der Signalverarbeitung. Sie geben in vielen Detailfragen (Trillerschwelle und Gesetz der Nähe)

schon Aufklärung darüber, was man überhaupt hört und wie man es hört. Die Schwebung, das Verschmelzen von Instrumentalklangfarben im Orchester, das Vibrato, ob es mehr als schwankende Tonhöhe oder als pulsierende Lautstärke gehört wird, findet hier seine Erklärung. Letztlich geht es um so entscheidende Fragen wie die nach den Grundlagen von Konsonanz, für die hier schon die Weichen gestellt werden.

Die Basilarmembran wird normalerweise als Filterbank mit ziemlich breiten Filtern vorgestellt. 24 von ihnen passen nebeneinander in den Hörbereich von unten bis oben, wenn man sie mit festen Grenzen nebeneinander legen würde. Oberhalb von c“ haben sie eine Breite von etwa einer Terz, dabei nach beiden Seiten ausladende Flanken, so dass man von genauen Grenzen kaum reden kann. Als Richtwert geht man von einer Intervallbreite von  $f_{go}/f_{gu} = 1,22$  aus, was etwa einer neutralen Terz entspricht. Unterhalb von 500 Hz haben die Frequenzgruppen eine Breite von etwa 100 Hz (Kuttruff 2004, S. 240). Innerhalb dieser Frequenzgruppen (critical bands), die durch diese Intervallbereiche beschrieben werden, liegt eine Kopplung vor, d. h. es findet eine gegenseitige Beeinflussung der Schwingungen statt. Deshalb ist auch der Ausdruck „Kopplungsbreiten“ geläufig. Die Positionen dieser Filterbereiche sind in Wirklichkeit aber variabel, sie passen sich dem jeweils eintreffenden Schall an und gruppieren sich um die analysierten Schallkomponenten ständig neu.

Der Filterwirkung entsprechend lösen bestimmte Frequenzanteile der Welle, die im Cochleakanal entlang läuft, an bestimmten Stellen der Basilarmembran maximale Bewegungen aus (Einortstheorie). Erste Eindrücke von den Bewegungsabläufen und der sehr groben Selektivität der Schallanalyse auf der Basilarmembran gaben uns die Bilder, die G. v. Békésy 1943 und 1947 veröffentlichte. Diese Bewegungen lösen in den Ganglienzellen Nervenimpulse aus, und zwar im Takt der jeweiligen Bewegungen, die an der von den Ganglienzellen jeweils abgetasteten Stelle vorliegen. Die Bewegungsausschläge werden getriggert.

Drei Dinge werden bei diesem Modell meist nicht beachtet:

1. Ein Formproblem, das die Filterkurve betrifft und sich auf die jeweilige schwingende Form der Basilarmembran auswirkt: Es handelt sich bei den Filtern um eine Serie von Tiefpässen mit sukzessiv abnehmender oberer Grenzfrequenz, nicht, wie meist in der Literatur zu finden, um Bandpässe.
2. Ein Zeitproblem, das durch die begrenzte Geschwindigkeit der Wanderwelle auf der Basilarmembran entsteht. Diese Filter werden von den unten in die Schnecke eintretenden Schwingungen (sie werden am ovalen Fenster von der Steigbügel-Fußplatte in das Innenohr hineingepumpt) nacheinander durchlaufen. Das bedeutet, die Welle kommt an den Stellen für die Registrierung tiefer Schallanteile später an. Die Welle gebraucht Zeit, um von der Basis bis zur Schneckenspitze zu gelangen.

Die Verzögerungszeiten für die Autokorrelation werden daher nicht nur durch Laufzeiten der Impulse auf den Nervenbahnen, sondern auch durch die Laufzeiten der Welle auf der Basilarmembran bereitgestellt.

3. Aus (1) und (2) folgt, dass aus dem Spektrum der Frequenzen von den höheren zu den tieferen fortlaufend mehr Schwingungsinhalt zurückgehalten wird, bis am Ende (in der Schneckenspitze) nur noch tieffrequenter Schallinhalt übrig bleibt. Mit dieser Art von Filterung ist verbunden, dass der tieffrequente Schallanteil an den Stellen, wo der höherfrequente Schallanteil abgetastet wird, entlang läuft und dort möglicherweise tieffrequente, seinem Schallinhalt entsprechende Triggerung auslöst. Dabei kann es zur Störung des Abtastvorgangs der höheren Frequenzen kommen (was sich durch eine stärkere Verdeckung höherer Schallanteile durch tiefere mehr auswirkt als umgekehrt; Stumpf, 1890, S. 220, 228 und 419).

Folge dieser Art der Schallanalyse ist also:

Auch an den Stellen, die für die Auswertung hoher Frequenzanteile vorgesehen sind, ist also die Abtastung tieffrequenten Schalls möglich; und sie erfolgt dort früher als an der Stelle, wo der tieffrequente Schall sein Maximum hat. Aber auch der umgekehrte Fall ist anzutreffen, wie wir sehen werden.

Die folgende Fallunterscheidung macht deutlich, was sich auf der Basilar-membran (BM) wirklich abspielt. Die dargestellten Vorgänge befinden sich in voller Übereinstimmung mit den neuesten experimentellen Ergebnissen:

Fall A. Wenn an einer Stelle der BM einige stärkere Teiltöne (TT) enger benachbart liegen als eine kritische Bandbreite, das ist oberhalb c“ etwa eine Terz, pulsieren sie mit der Frequenz ihrer Differenz. Sie bilden sozusagen eine Schwebung; diese hat bei harmonischen TT die Frequenz des Residuums. Die Gruppe aus zwei oder mehr TT signalisiert dem Gehörorgan folgendes:

1. Der Ort der Reizung hat die Bedeutung: es liegt eine diesem Ort entsprechende klangfarbliche Auswertung vor: Formantwirkung (tonotope Abbildung).
2. Grundtonfrequenz ist gegeben durch die Frequenz des Pulsierens der BM an dieser Stelle (periodotope Abbildung). Das heißt: Auswertung des Residuums findet an einer Stelle statt, die der klangfarblichen Auswertung höherer Frequenzen zugeordnet ist.

Fall B. Wenn an einer Stelle ein einzelner Teilton stärker hervortritt, dessen benachbarte TT mehr als eine Quarte Abstand von ihm haben, vibriert die Stelle mit seiner Frequenz und liegt auch an der dieser Frequenz zugeordneten Stelle. Sie signalisiert:

1. Aus der an diesem Ort vorliegenden Reizung folgt die klangfarbliche Wirkung dieser Komponente (tonotope Abbildung).
2. Die Bewegungsfrequenz an dieser Stelle ist die der Spektralkomponente (Der Reizort für die periodotope Abbildung ist der gleiche wie der für die tonotope).

Aus diesem Sachverhalt ist folgender Schluss zu ziehen:

1. Klangfarbe wird vermittelt durch den Ort auf der BM, und zwar unabhängig davon, mit welcher Frequenz diese Stelle vibriert,
2. Residuum kann an jeder Stelle der BM abgeleitet werden.
3. Verzögerungszeiten kommen auch durch die Ausbreitung der Welle auf der BM zustande.

Es handelt sich hier um einen viel zu wenig beachteten Sachverhalt, der durch experimentelle Studien aber belegt und aus ihren Ergebnissen zu erschließen ist, insbesondere aus denen von Langner et al. (1988, 2000), Langner (1995) und Tramo et al. (2001). Die dargestellten Vorgänge spielen bei der Formantwirkung der Klänge und bei der Konsonanzempfindung eine entscheidende Rolle.

### *2.3 Neuronale Konfiguration (Abbildung im Cortex)*

Durch die Hirnforschung der letzten 100 Jahre sind folgende Tatsachen als gesicherte Ergebnisse anzusehen:

1. Den Sinnesorganen sind in der Großhirnrinde (cortex) bestimmte Areale zugeordnet.
2. Innerhalb eines Sinnesgebietes, hier speziell des Hörens, aktivieren benachbarte Rezeptoren sogar benachbarte Positionen innerhalb der Cortex-Areale. Es gilt das Gesetz der Kontinuität und der Nachbarschaft für ähnliche Sinnesreize. Benachbarte Sinneszellen, die im Sinnesorgan aktiviert werden, haben eine Aktivität benachbarter Stellen in der Großhirnrinde zur Folge. Es findet also eine regelrechte Abbildung im Cortex statt. Sie kann beschrieben werden durch eine lineare Transformation.

Für Töne verschiedener Tonhöhe gibt es auf der BM Stellen maximaler Erregung der Nerven in Abhängigkeit von der Frequenz dieser Töne (Einortstheorie). Hohe Töne erregen sie maximal an der Basis des schneckenförmigen Innenohres, tiefe Töne an der Spitze. Durch die Erregungsmuster im Cortex konnte diese Anordnung auf der BM, die sich auch in den Hörnerven widerspiegelt, bestätigt werden. Langner und Schreiner (1988, S. 1799–1822, Fig. 3, S. 1804) haben orthogonal zu dieser Anordnung auch die Frequenzen der periodotopen neuronalen Auswertung abgebildet gefunden. Dementsprechend sind benachbarte durch Periodizität gegebene Tonhöhen auch in einer Reihe benachbarter Punkte auf der Großhirnrinde abgebildet.

Wenn diese Art von Abbildung bis in die Endstation zu verfolgen ist, dann ist sie auch auf den Zwischenstationen der neuronalen Verarbeitung zu vermuten. Langner et al. (1988, 2000) sowie Schreiner et al. (1988) haben daher in einer wichtigen Schaltstation, dem Colliculus Inferior (IC) und darin im

Zentrum (ICC), Untersuchungen über die Anordnung der Tonhöhen vorgenommen und dabei speziell danach gesucht, wie Periodizität ausgewertet wird. Sie arbeiteten daher nicht nur mit Sinustönen, sondern mit modulierten Tönen (sinusförmige Modulation 100 % und Pulstönen), die eine starke Periodizität aufweisen. Sie stellten fest, dass diese Periodizität an einer bestimmten Stelle durch erhöhte Aktivität der Nerven nachweisbar ist und diese Stelle nicht mit der Stelle der maximalen Erregung auf der BM (der tonotopen Abbildung) identisch ist. Mit anderen Modulationsfrequenzen gab es auch andere Abbildungen der Periodizität, und zwar wurden auch diese in einer Reihe angeordnet gefunden. Die Achsen dieser Anordnung waren (etwa) orthogonal. Das Ergebnis also war eine „periodotope Abbildung“ mit einer Darstellung der „best frequency“, die wie eine Resonanzkurve aussieht, genau genommen aber angibt, wie stark eine bestimmte Nervenfasern auf verschiedene Frequenzen reagiert. Man sagt auch, wie scharf sie auf „ihre“ Frequenz abgestimmt ist. Die hiermit bewiesene periodotope Abbildung ist unabhängig von der tonotopen (Langner, 1995; Langner et al., 1997, 2000).

Damit war die lange vermutete und lange gesuchte Auswertung der Periode in der Schallanalyse des Gehörs auch neurophysiologisch nachgewiesen. Die Tatsache, dass die Nervenimpulse von der Modulationsbewegung synchronisiert werden, die durch die (mindestens drei) Spektralkomponenten der Modulation hervorgerufen wird, lässt den Schluss zu, dass jede Gruppe genügend nah benachbarter Spektralkomponenten, also auch eine Gruppe benachbarter harmonischer Teiltöne eines Klangspektrums, eine solche Synchronisation hervorrufen wird. Modulierte Töne, die spektral betrachtet aus einer Trägerfrequenz mit benachbarten Seitenbändern bestehen, verursachen auf der BM an der Stelle der Trägerfrequenz und im angrenzenden Gebiet oberhalb und unterhalb, aber jedenfalls in einem Kopplungsbereich (= critical band), maximale Bewegung und damit auch maximale Erregung der Ganglienzellen. Ihre Pulsationsfrequenz ist dort aber wie bei einer Schwebung die des Unterschieds der Frequenzen; denn sie ist durch den Abstand der Spektralkomponenten gegeben. Sie würde aber, wenn sie als Frequenzkomponente real im Schall vorhanden wäre, an einer Stelle für die Abtastung viel tieferer Frequenzen liegen. Bei harmonischen Komponenten ist diese Differenz zwischen zwei benachbarten Teiltönen grundsätzlich gleich dem größten gemeinsamen Teiler (g. g. T.), und wir hören auf dieser Frequenz den Residualton. Deshalb hat die Entdeckung der neuronalen Periodizitätsauswertung auch so hohen Erkenntniswert. Sie steht seit A. Seebeck (1841) und der nachfolgenden Ohm-Seebeck-Kontroverse (1843, 1844) in der Diskussion (Schouten, 1940a, 1971), ohne dass bisher eine schlüssige Erklärung hätte gegeben werden können.

Stellt man sich diese Vorgänge einmal bildlich anhand eines tiefen Fagotttones vor, den man aus der Entfernung von einigen Metern hört, dann dürfte man von folgendem Erregungsmuster auf der BM ausgehen:

Der Grundton des Tones C ist so schwach, dass er nicht gehört wird. Dies entspricht dem Grundsatz, dass alle Musikinstrumente zu klein sind, um die Grundtöne ihrer tiefen Töne kräftig abstrahlen zu können. Bei einem Fagott

beträgt die Schalldruckamplitude des Grundtones durchschnittlich nur etwa 10 % der kräftigen Spektralanteile, was einem dB-Wert von  $-20$  dB entspricht. In Kombination mit der Tatsache, dass die Hörschwelle bei einer Frequenz von 65 Hz für den Ton C in der Regel schon 30 dB höher liegt als bei 1 kHz ( $c^3$ ), ist davon auszugehen, dass die Stärke der Grundtonkomponente dann nicht ausreicht, um über die Hörschwelle zu treten. Seine weitere Berücksichtigung kann in den folgenden Überlegungen daher vernachlässigt werden.

Die kräftigen Spektralanteile findet man in den beiden Formantbereichen, nämlich beim 6. bis 8. Teilton im Bereich der Tonhöhe um  $a^1$  (450 Hz) und bei den Teiltönen 15 bis 19 im Bereich der Tonhöhe  $d^3$  (Mertens 1975, S. 57; Reuter 2002, S. 264). Beide Abschnitte auf der BM, in denen diese Frequenzen nach der Einortstheorie ihre Positionen haben, werden kräftige Bewegungen mit starken Ausbauchungen der BM ausführen. Diese Bewegungen sind aber nicht gleichförmig, sondern durch ein starkes Pulsieren gekennzeichnet. Die Frequenz des An- und Abschwellens wird in beiden Fällen nicht die sein, die eine einzelne dort befindliche Spektralkomponente haben würde, sondern aufgrund des Zusammenwirkens mehrerer Teiltöne in einer Kopplungsbreite wird es die des Grundtones sein. Schon je zwei benachbarte Teiltöne würden aufgrund des Schwebungseffekts Schwebungsbewegungen mit der Frequenz ihrer Differenz, d. h. im harmonischen Falle der des Grundtones, zeigen. Nur in der Feinstruktur der Schwebung ist die zum Abbildungsort gehörige resultierende Frequenz enthalten. Bedingung für derartige Erscheinungen ist in jedem Fall die fehlende oder die mangelnde Analysierschärfe, die hier für Teiltöne oberhalb vom 5. Teilton gegeben ist.

Festzuhalten ist also: wenn zwei oder mehr Harmonische eines Formantbereichs in eine Frequenzgruppe fallen, führen diese Formantbereiche Bewegungen mit der Frequenz des Grundtones aus. Folglich werden von dieser Stelle Nervenimpulse ausgesandt, die mit der Frequenz des Grundtones synchronisiert sind und nicht solche mit der für diese Stelle vorgesehenen Frequenz. Genau das haben Langner et al. (1988) mit ihren Modulationsversuchen herausgefunden. Bei Modulationen sind es immer mindestens drei Spektralkomponenten; werden sie so gewählt, dass sie in einem harmonischen Verhältnis zueinander stehen, ist der größte gemeinsame Teiler (g. g. T.) der Grundton, der als Residualton (innerhalb der gut recherchierten existence region (Walliser, 1968, 1969)) auch wahrgenommen wird.

## *2.4 Periodizität und größter gemeinsamer Teiler bei Intervallen*

Gleiche Tonkombinationen wie in der Obertonreihe findet man unter den harmonischen Intervallen. Die Intervalle des harmonischen Tonsystems der abendländischen Musik bilden eine Auswahl aus dieser Reihe, wobei eine Grenze zwischen dem 6. und dem 7. Teilton gezogen ist, um die Zahl 7 aus unserem System auszuschließen. (Seit der Zeit des Pythagoras bis Didymos lag die Grenze noch bei der Zahl 4. Die Pythagoräer rechneten nur mit der Tetraktys, der Gruppe der ersten 4 Zahlen. Alle Intervalle unseres heutigen



Es fehlte dieser Theorie der Konsonanz (eigentlich) nur der anatomische und neurophysiologische Hintergrund, um sie endgültig zu etablieren. Darüber hinaus fehlte ihr die Erklärung, warum sich diese Regelmäßigkeit des Zusammentreffens im Hörvorgang positiv auswirkt. Die Regelmäßigkeit mathematisch zu erfassen, gelingt mit der Autokorrelation. Licklider (1951) und Meyer-Eppler (1951, 1959) hatten die Autokorrelation schon zur Ermittlung periodischer Vorgänge in Schallgemischen vorgeschlagen. Zwei Dinge irritierten aber damals und irritieren bis heute bei dieser mathematischen Formulierung:

$$a_t(\tau) = \int_{-\infty}^{\infty} f(t) \cdot f(t + \tau) dt$$

**Formel 1:**

Formel für die Autokorrelation

Sinusförmige Komponenten des Schalls  $f(t) = \sin \omega t$ , aus denen der Schall nach Fourier zusammengesetzt ist, ergeben nach dieser Formel nur schwach von einander abgesetzte Maxima und damit eine schlechte Trennung der einzelnen Komponenten. Setzt man in die Autokorrelationsanalyse anstelle der Sinuskomponenten Nadelimpulse oder Dirac-Stöße ein, wird die Trennung so scharf, dass sie der Wirklichkeit neuronaler Verarbeitung nicht mehr entspricht. Sie liefert diskrete Frequenzpunkte bei genauer Koinzidenz, ohne Übergänge und Zwischenwerte.

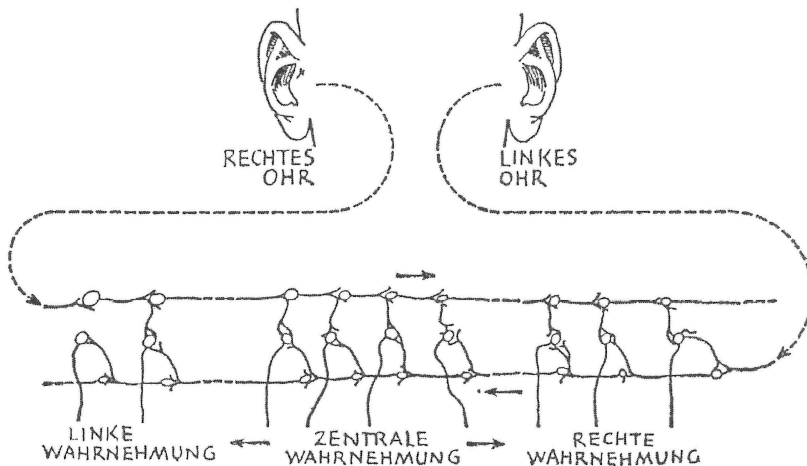
## 2.5 Neuronale Vernetzung und Gleichzeitigkeit

Man kennt inzwischen ziemlich genau die Vernetzungsstruktur in der Großhirnrinde und auf den verschiedenen Zwischenstationen von den Sinnesorganen bis dorthin. Eine der Schaltstellen, die sich für das Hören als entscheidend herausgestellt haben, ist der Colliculus inferior (CI) und darin das Zentrum (ICC). Von den verschiedenen Verschaltungsmöglichkeiten interessieren hier zunächst die, bei denen Dendriten von mindestens zwei präsynaptischen Neuronen an einer Nervenzelle anliegen und diese postsynaptische Zelle dadurch veranlasst wird, einen Impuls auf ihrem Axon loszuschicken, wenn mindestens zwei Impulse gleichzeitig von den präsynaptischen Neuronen ankommen. Nur beide Impulse zusammen heben das Potenzial in der postsynaptischen Zelle so weit an, dass diese selbst einen Impuls aussendet. Diese vereinfachende Darstellung soll nur das Prinzip der Autokorrelation beschreiben. In Wirklichkeit empfängt jede postsynaptische Zelle von zahlreichen (bis zu 10.000) Synapsen anderer Zellen Signale und sendet Signale an ebenso viele andere Zellen. Erst die von zahlreichen Synapsen zusammen ausgehenden Impulse heben das Potenzial der Zelle über die Schwelle, die zum Feuern notwendig ist. Für eine angemessene Behandlung dieser Vor-



gänge bietet sich eine statistische Betrachtung der neuronalen Prozesse an (Fricke, 2009).

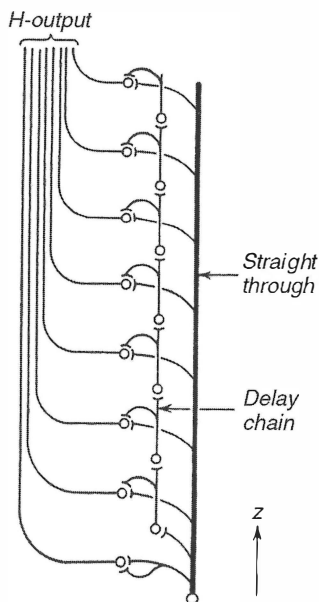
Ein solcher Mechanismus ist in der Lage, Gleichzeitigkeit festzustellen wie ein logisches Schaltelement der Elektronik, das als Und-Gatter bezeichnet wird. Solche Nervenschaltungen vermutete man schon lange zur Erklärung des Richtungshörens, und zwar jenes Teils zur Richtungsfeststellung, der auf der Zeitdifferenz zwischen den beiden Ohren beruht (vgl. Abb. 2 und 3; Fig. 8 von Licklider, 1959, S. 100 und Bergeijk et al., 1960 nach den Vorschlägen von Jeffress, 1948). Hier werden Laufzeiten auf den Nervenleitungen und die Feststellung der Gleichzeitigkeit des Eintreffens zweier Impulse benötigt, um jeder Zeitdifferenz zwischen den beiden Ohren die Aktivierung einer bestimmten Zelle bzw. Zellgruppe zuzuordnen. Zeitunterschiede werden damit in Ortskoordinaten transformiert. Es ergibt sich hier sozusagen eine Ortstheorie des Richtungseindrucks.



**Abb. 2:**

Koinzidenz von unterschiedlich verzögerten Nervenimpulsen zur Richtungserkennung (nach Bergeijk et al., 1960, S. 167)

Die Kombination von Verzögerungszeiten, die man durch Umweschaltungen und Verzögerungselemente erreichen kann, mit Und-Gattern, in denen die Gleichzeitigkeit festgestellt wird, bietet sich also grundsätzlich an, um Zeitunterschiede oder Zeitstrecken zu messen. Bei Tönen besteht so die Möglichkeit, die Dauern ihrer Perioden  $T$  und damit (indirekt) ihre Tonhöhe zu ermitteln. Würde dieses Verfahren zur Messung der Frequenz  $f = 1/T$  eingesetzt, würde gleichzeitig „the case of the missing fundamental“ (Stevens & Davis, 1938, S. 99; nach Fletcher, 1934) gelöst. Denn die Periode findet sich in jeder Gruppe benachbarter Obertöne wieder, die die Basilarmembran in Bewegung setzen.

**Abb. 3:**

Lickliders Darstellung eines neuronalen Autokorrektors auf der Basis einer Koinzidenz unterschiedlich verzögerter Nervenimpulse zur Messung der Periodendauer (Licklider, 1959, S. 102)

Als Alternative gibt es – wenn es gilt Frequenzen zu messen – nur die Möglichkeit, die Zahl der Perioden pro Zeiteinheit zu zählen. Technische Frequenzzähler bieten beide genannten Möglichkeiten an, damit man die bequemere unter ihnen wählen kann. Eine dritte Möglichkeit ist die, Frequenz durch Resonanz zu ermitteln, wie dies bei einem Klavier mit aufgehobener Dämpfung oder mit einer Harfe möglich wäre. Helmholtz hatte dieses Bild vor Augen, als er seine Resonanztheorie des Hörens entwarf. Die Natur hat sich für diese Lösung wohl aus technischen Gründen nicht entschieden, sondern statt dessen den Schneckenkanal mit Trennwand entwickelt, auf der eine Schlauchwelle entlang läuft. Diese Welle durchläuft zwar auch ein Maximum, das in Abhängigkeit von der Frequenz an verschiedenen Stellen des Schneckenkanals liegt (Einortstheorie). Die Abbildung der Frequenz auf der Längsachse der Basilarmembran ist jedoch so grob, dass sie nur der Vorsortierung der Spektralanteile verschiedener Frequenzen dient und zur Erklärung von Schwebungen, Modulationen, Lautheitsaddition und Verschmelzungserscheinungen herangezogen werden kann.

Im Vergleich zu den Schaltelementen der Elektronik, die praktisch unendlich schnell reagieren, ist die Reaktion in Nervenschaltungen langsamer. Die Nervenimpulse haben selbst eine endliche Dauer, und die Reaktionszeit der Zelle, die auf die Impulse reagiert, ist auch nicht unendlich schnell

(Dudel, 1990, S. 24 u. 52; Koch, 1999, S. 142). Das elektrische Potenzial in der Zelle muss durch die innerhalb einer Integrationszeit eintreffenden Impulse erst angehoben werden. Und erst, wenn es eine Schwelle überschreitet, reagiert die Zelle mit einem Impuls. Es gibt also zwei Ursachen für die endliche Breite der Reaktionszeit (Fricke, 2005a, S. 134–135).

Dies hat entscheidende Folgen für die Analysierschärfe der Autokorrelation. Die mathematische Autokorrelation von Impulsfolgen, wie sie die Zeitreihen-Autokorrelation nahelegt, liefert nur diskrete Werte für die Frequenzen. Bei Intervallen, deren Frequenzverhältnis durch einfache ganze Zahlen definiert ist, zeigt die Autokorrelation deshalb nur dann die Koinzidenz zweier Impulse an, wenn diese genau zusammentreffen.

$$\int_{-\infty}^{\infty} \delta(t-t_1) \cdot \delta(t-t_2) dt = \delta(t_1-t_2)$$

**Formel 2:**

Das Produkt zweier Dirac-delta-Funktionen (Nadelimpulse) ergibt wieder eine Dirac-delta-Funktion (nach Hartmann, <sup>2</sup>1998, S. 155, Gl. 7.22)

Diese mathematische Autokorrelation ist ein Pendant zur elektronischen Schaltung mit einem logischen Schaltelement. Es reagiert nicht, wenn Abweichungen vorhanden sind, seien sie auch noch so gering. Aus diesem Grund vermag das Modell der Autokorrelation bzw. der Zeitreihenanalyse die musikalische Konsonanz nicht ausreichend abzubilden. Das Modell ist hinsichtlich der Verstimmungstoleranz zu restriktiv. Die Schärfe des Zusammentreffens, eine Genauigkeit, die durch die Mathematik in die Modellrechnung hineingetragen wird, indem man mit idealisierten Impulsen anstelle von realen rechnet, wird erst durch das „Zeitfenster der Gleichzeitigkeit“ beseitigt.

## *2.6 Die ISI-Ergebnisse von Tramo et al. als ein Beweis für Autokorrelation mit Unschärfe*

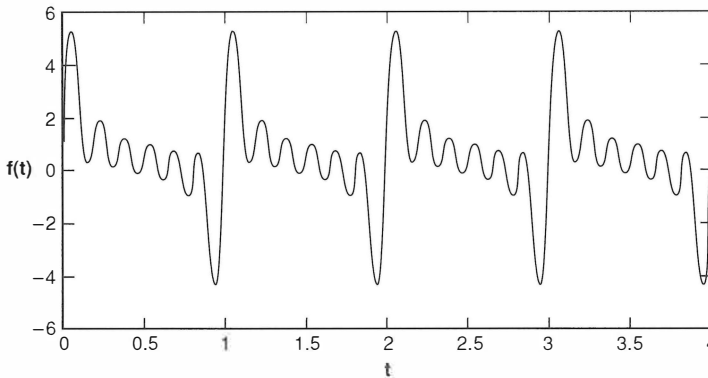
Tramo et al. (2001) registrierten Interspike-Intervalle (ISI), die sie vom Hörnerv von Katzen ableiteten. Sie erfassten mit der Ableitung eine größere Zahl von Nervenfasern (> 100) und registrierten die Abstände der Nervenimpulse (Spikes) in Histogrammen. Die Histogrammverteilungen zeigen nicht nur die Abstände zwischen je zwei benachbarten (first order), sondern auch die zwischen weiter auseinander liegenden Spikes. Bei der Darbietung der Töne verwendeten sie konsonante Intervalle (Quinte und Quarte in den Frequenzverhältnissen 3/2 und 4/3) und dissonante Intervalle (Ganzton und Tritonus in den Verhältnissen 9/8 und 45/32). Unter den dissonanten Intervallen stellte sich aber das Intervall 45/32 nicht als so dissonant dar, wie man es aufgrund

seiner Zahlen hätte erwarten dürfen. Denn sowohl die neuronale Auswertung als auch die Simulation per Modellrechnung stufen es als verstimmtes, schwach konsonantes Intervall 7/5 ein. Das ist verständlich, wenn man folgende zwei Dinge bedenkt:

1. Die Autokorrelation wurde mit einer unscharfen Koinzidenzprüfung durchgeführt. Diese Unschärfe der Autokorrelation ist sowohl in der Simulation versteckt als auch in den gemessenen ISIs zwangsläufig enthalten. So konnte 45/32 als 7/5 gewertet werden.
2. Das Intervall 7/5 kommt in dieser Form in unserem abendländischen Ton-system zwar nicht vor, weil die Zahl 7 ausgeschlossen ist. In der Berechnung der Autokorrelation aber steht es im Rang sogar vor der kleinen Sexte 8/5, ist also nicht als so scharfe Dissonanz zu werten wie der Tritonus 45/32. Dies ist der Grund, weshalb es sowohl in den ISIs als auch in der Simulation als schwach konsonant eingestuft wurde.

Zu (1) ist zu erläutern, dass verstimmte Konsonanzen, die noch konsonant klingen, mit einer Autokorrelationsanalyse dann und nur dann erkannt werden können, wenn die endliche Impulsbreite und die Unschärfe der Gleichzeitigkeit in der Autokorrelation berücksichtigt werden (Fricke, 2005a, S. 134–135). Autokorrelation misst die Koinzidenz der Impulse. In der Zeitreihen-Autokorrelation (Hesse, 2003, S. 142) von Punktreihen, wie sie Euler (1739, Tabelle I zwischen S. 36 und 37) veranschaulicht, wird deutlich, dass diese Punkte eine Idealisierung darstellen, d. h. eine Abstraktion von den Schwingungsperioden oder Nervenimpulsen. In diesem idealisierten Fall wird eine Koinzidenz nur festgestellt, wenn das Intervallverhältnis mathematisch genau auftritt (s. o. Formel 2). Die sogenannte neuronale „Fehl-messung“ bei Tramo ist also ein Indiz dafür, dass bei der neuronalen Verarbeitung im Gehör der Katze auch unscharfe Intervallverhältnisse registriert werden. Sie beruht darauf, dass Nervenimpulse keine idealen Impulse, keine Nadelimpulse oder Dirac-Stöße sind, sondern eine gewisse Breite haben, die in der Größenordnung von Mikrosekunden liegt (Dudel, 1990, S. 24). Auch wird das gleichzeitige Eintreffen mehrerer Impulse an der postsynaptischen Zelle, die die Funktion eines Und-Gatters hat, diese erst zum Feuern veranlassen, wenn das Potenzial in ihr das erforderliche Niveau erreicht hat (Dudel, 1990, S. 52). Beides benötigt Zeit und hat eine „Unschärfe der Gleichzeitigkeit“ zur Folge. Die „Fehl-messung“ in der Simulation hingegen beruht offensichtlich auf der Breite der Impulse, die dabei eingesetzt wurden. Diese Breite kommt dadurch zustande, dass die Simulation der Impulse mit nur 6 Harmonischen vorgenommen wurde. Wie Ebelings Synthese zeigt (Abb. 4), ergeben sich bei der Superposition von 6 Harmonischen gleicher Amplitude am Anfang jeder Periode Impulse endlicher Breite. Die Zusammensetzung der Impulse erfolgte hier mit 6 Sinusschwingungen, bei Tramo et al. (2001) dagegen mit 6 Cosinusschwingungen gleicher Amplitude. Der Unterschied liegt nur in den Phasen des Anfangs jeder Periode, der aber bei der Autokorrelation keine Rolle spielt, da die Autokorrelationsfunktion als ein Analysever-

fahren gilt, das phasenunabhängig arbeitet. (Die den Autokorrelogrammen erstaunlicherweise sehr ähnlichen Schwingungsbilder, die bei Tramo et al. (2001) gezeigt werden, sind ebenfalls auf diese besonderen Synthesebedingungen der Intervalltöne zurückzuführen). Dies ist ein typisches Beispiel dafür, wie eine spezielle Modellbildung die nachfolgenden Ergebnisse vorwegnimmt.



**Abb. 4:**

Darstellung der aus sechs Harmonischen zusammengesetzten Impulse (nach Ebeling, persönliche Mitteilung), wie sie von Tramo et al. (2001) zur Berechnung der Autokorrelation der aus solchen Tönen gebildeten Intervalle verwendet wurden. Die Schwingung zeigt deutlich die endliche Breite der Impulse.

Zu (2): Der von Tramo et al. (2001) unter den dissonanten Intervallen genannte Tritonus wurde mit dem Schwingungszahlverhältnis  $45/32$  modelliert. Dieses Intervall gehört dem harmonischen System an und ist darin als übermäßige Quarte zu interpretieren. Man erhält es, wenn man von der Terz  $5/4$  einen Ganzton  $9/8$  herauf geht oder von der Quinte  $3/2$  einen Halbton  $16/15$  abwärts schreitet. Dieses Intervall liegt in unmittelbarer Nachbarschaft des aus der Naturseptime gebildeten Tritonus  $7/5$  im Abstand von 8 Cents (Hundertstel temperierter Halbton). Dass die Naturseptime in unserem abendländischen Tonsystem keine Verwendung findet, hat das Analyseverfahren nicht daran gehindert, das Intervall  $7/5$  aufzuspüren. Von der Autokorrelation mit unscharfer Koinzidenz wurde es sowohl im Fall der ISI-Registrierung als auch im Fall der Simulation als noch relativ einfaches Intervall registriert. Dies ist ein erster Hinweis darauf, dass in der neurophysiologischen Auswertung auch verstimmte Intervalle richtig erkannt werden.

Tramos und Ebelings Berechnungen erbrachten also nur deshalb den Beweis, dass auch verstimmte Konsonanzen von der Autokorrelation erkannt werden, weil sie mit nur 6 Harmonischen rechneten und auf diesem Wege eine durch die Breite der Impulse zustande kommende Unschärfe simulierten. Tramos ISI-Registrierungen sind nebenbei ein Hinweis darauf, dass schon vor dem Nucleus Cochlearis eine Intervallerkennung stattfindet. Durch die

neurophysiologische Auswertung werden dort sogar schon verstimmte Intervalle richtig erkannt. Beides ist offensichtlich dadurch nachgewiesen, dass die Ableitung für die Registrierung am Hörnerv erfolgte und dort die Autokorrelationsprodukte schon vorlagen.

Fasst man die Forschungsergebnisse der letzten Jahre zusammen, ergibt sich folgendes Bild:

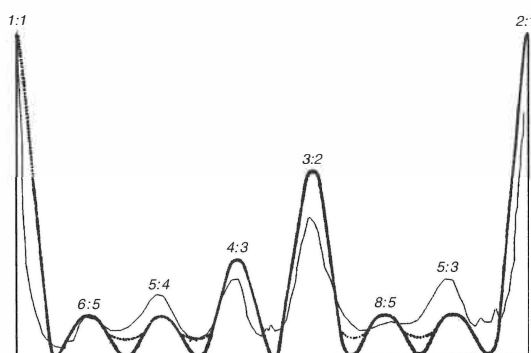
Verstimmte Konsonanzen, die aber dennoch konsonant klingen, können mit einer Autokorrelationsanalyse erkannt werden, wenn die endliche Impulsbreite und die Unschärfe der Gleichzeitigkeit in der Autokorrelation berücksichtigt werden (Fricke, 2005a, S. 134–135).

Dieses Zeitfenster der Unschärfe kann auf zweierlei Weise modelliert werden. Entweder man bildet ein Rechteckfenster für die Zeit, in der zwei Nadelimpulse dann als „gleichzeitig“ erkannt werden, wenn sie innerhalb dieses Fensters liegen, oder man bildet anstelle der Nadelimpulse schmale Rechteckimpulse, die dann als gleichzeitig gelten, wenn sie sich zeitlich überdecken. Die Summe der beiden Rechteckimpulse in diesem Überlappungsbereich (das „Überlappungsintegral“ nach Ebeling) ergibt ein Dreieck, dessen Basisbreite gleich der Summe zweier Rechteckimpulse ist.

Beide Formen der Unschärfedarstellung sind neurophysiologisch existent:

1. Die Breite der Nervenimpulse liegt im Mikrosekundenbereich und kann mit 0,05 bis 1,0 ms angenommen werden.
2. Für die Reaktionszeit eines Neurons sind Latenzzeiten von  $1/4$  bis  $1/3$  ms bekannt. Wenn innerhalb dieser Zeit Impulse von präsynaptischen Neuronen ankommen, wird Gleichzeitigkeit signalisiert und das postsynaptische Neuron aktiviert; es feuert.

Beide „Zeitbreiten“ wirken also zusammen. Sie werden im Modell am besten durch den Wert 0,8 ms simuliert. In diesem Fall liefert die Rechnung eine



**Abb. 5:**

Verschmelzungsgrade harmonischer Intervalle nach Stumpf (1890, S. 176, fette schwarze Linie) im Vergleich zur berechneten Autokorrelation mit Unschärfe nach Ebeling (persönliche Mitteilung, dünne Kurve) für die Intervalle zwischen Prim und Oktave.

Kurve, die der von Stumpf (1890, S. 176) mitgeteilten Kurve über die Verschmelzungsstufen konsonanter Intervalle sehr ähnlich ist (vgl. Abb. 5).

In der durch Korrelation berechneten Kurve befinden sich außerdem noch bei  $7:5$  und  $10:7$ , d. h. in den beiden Versionen des durch die Naturseptime gebildeten Tritonus (übermäßige Quarte und verminderte Quinte), schwach erkennbare Erhebungen. Diese „Konsonanzen“, von denen die mit der Proportion  $7:5$  sogar „einfacher“ ist als die der kleinen Sexte  $8:5$ , konnte Stumpf aus zwei Gründen nicht würdigen.

1. Die Zahl Sieben passt nicht in die Vorstellungswelt der damaligen Musiktheorie. Intervallverhältnisse auf der Basis Sieben waren von den Tonsystemen ausgeschlossen.
2. In dem vorgefertigten Tonmaterial, das Stumpf damals auf Klavieren und Orgeln zur Verfügung stand, waren diese Intervalle deshalb nicht enthalten.

Die in den Lehrbüchern (Meyer-Eppler, 1959, S. 10) mitgeteilte Autokorrelationsfunktion zur Durchmusterung von Schallvorgängen, um periodische Anteile in ihnen aufzuspüren, liefert im Ergebnis nur weiche Übergänge von einem Maximum zum anderen. Die Autokorrelation von Punktmustern, wie z. B. die oben genannten bei Euler (1739, Tabelle I zwischen p. 36 und p. 37), die eine Digitalisierung des Schwingungsverlaufs darstellen, zeigt im Ergebnis nur diskrete Punkte anstelle der runden Maxima. Im Falle des o. g. Intervalls  $45/32$  würde das bedeuten, dass ein phasengleiches Zusammenreffen (Koinzidenz) der beiden Schwingungen erst nach 45 Schwingungen des höheren mit 32 Schwingungen des tieferen Tones zustande kommt. Abzulesen ist aus dem bei Tramo et al. wiedergegebenen Diagramm (2001, S. 100) aber, dass dies nach 11,4 ms schon der Fall ist. Dies beweist, dass die Nervenschaltung bei der Autokorrelation ein Zeitfenster für die Unschärfe benutzt und dadurch in der Lage ist, konsonante Intervalle, auch wenn sie verstimmt sind, als einfache Intervallverhältnisse zu erkennen. Ein der Realität am ehesten entsprechendes Modell muss also zwischen diesen Extremen analoger und digitaler Durchmusterung positioniert sein und wird z. B. durch die Berücksichtigung des o. g. Zeitfensters von 0,8 ms gewonnen.

### 3 Ein Modell simuliert ein Modell

Stumpf hatte sehr ausführliche experimentelle Untersuchungen zur Verschmelzung durchgeführt. Sie erstreckten sich über einen langen Zeitraum und konzentrierten sich auf zwei Perioden, von denen die eine noch in Prag lag, die andere 1888 in Halle folgte. Im Grunde zielte der ganze, sehr umfangreiche zweite Band der „Tonpsychologie“ (1890) darauf ab, Stumpfs Verständnis von Verschmelzung zu erklären. Er beschreibt vor diesem Hintergrund die durchaus noch mögliche getrennte Wahrnehmung zweier Töne, die ein Intervall bilden, bei gleichzeitiger Vereinigung dieser Töne zu einer Einheit. Diese Einheit ist die Verschmelzung, die er am liebsten „Schmelz“ oder

„Schmalz“ (S. 129) genannt hätte, weil sie nach seiner Meinung eher einem Produkt als einem Prozess gleichzusetzen ist. Für diese „Verschmelzung“ genannte Einheit gibt es graduelle Unterschiede, die in einer breiten Skala von Fallunterscheidungen darzustellen sind. Die durchaus noch vorhandene Möglichkeit, zwei Töne noch als zwei Intervalltöne zu erkennen, auch wenn sie gleichzeitig erklingen, wie das bei zwei Klaviertönen beispielsweise noch der Fall ist, ist der weitaus geringeren Möglichkeit gegenüberzustellen, zwei obertonarme Orgeltöne (Aliquoten z. B.) noch als zwei Einzeltöne auseinander zu halten oder gar zwei Obertöne in einem Ton (akustisch „Klang“), der aus harmonischen Obertönen besteht, zu trennen. All dies wird aufgrund von eigenen und aus der Literatur bekannten Einzelbeobachtungen ausführlich erörtert und gegeneinander abgewogen.

Die verschiedenen Möglichkeiten, Töne noch getrennt wahrzunehmen, wenn sie gleichzeitig erklingen, sind aber nicht nur abhängig vom Klangmaterial, aus dem die Töne bestehen – ob sie einen harmonischen Obertonaufbau haben und welchen, ob sie unharmonische und geräuschhafte Beimischungen haben und welche – sie sind auch abhängig von der Begabung und dem Training des Hörers, von der momentanen Einstellung (analytisch – ganzheitlich) und schließlich vom Wichtigsten: dem Abstand der Töne und dabei insbesondere vom Intervallverhältnis, ob es harmonisch ist und wie genau es die harmonische Proportion einhält. Hier schließen sich sogleich die damals schon vorhandenen Erfahrungen mit der Verstimmungstoleranz an, die Stumpf bekannt waren und deshalb von ihm auch in der Kurve berücksichtigt wurden.

Bei aller Würdigung der experimentellen Basis, aus der Stumpfs Erkenntnisse hervorgegangen sind, darf doch nicht unbeachtet bleiben, dass er seine experimentellen Befunde nicht angemessen statistisch auswertete. Dass seine Ergebnisse nicht heutigem wissenschaftlichem Standard entsprechen, bedarf nicht der Erwähnung, hervorzuheben ist aber die Tatsache, dass seine Verschmelzungskurve eine von Hand gezeichnete Kurve ist, die (nur) seine Vorstellungen von Verschmelzung, die er in den 7 Verschmelzungsgesetzen formulierte, wiedergibt. Sie spiegelt die damals schon mindestens 600 Jahre alte Lehrmeinung der abendländischen Musiktheorie wieder, die vorherrschte, seit Franco von Köln die Terzen in den Kreis der Konsonanzen einbezog, und nennt die Konsonanzen bis zur Quarte in der Rangfolge, die der pythagoräischen Schule schon bekannt war. Ausdrücklich neu ist, wenn man von Äußerungen von Mengoli (1670, S. 56 ff.), Eximeno (1774) und Kiesewetter (1846) einmal absieht, die Berücksichtigung der Toleranzbreiten, die sich grafisch in der Breite der sich über den harmonischen Proportionen erhebenden Maxima ausdrückt.

Die Leistung Stumpfs besteht darin, mithilfe der Verschmelzung dasjenige Phänomen gefunden und benannt zu haben, das vorrangig für die musikalische Konsonanz verantwortlich ist, und gezeigt zu haben, dass es mit der von den Musiktheoretikern seit Jahrhunderten beschriebenen Konsonanz konform verläuft. Entscheidend ist dabei die Parallelität, die nachgewiesen wurde; denn sie stützt die Tatsache, dass mit der Verschmelzung eine zutreffende Eigenschaft für die Verbindung zweier Töne benannt wurde. Wenn



nun in einem zweiten Schritt durch die Koinzidenzeigenschaften neuronaler Spikes die Verschmelzung erklärt werden kann – und zwar mit allen ihren in der Praxis der Musik vorherrschenden Erscheinungen – dann ist das eine weitere Stütze für diese Theorie.

Ebelings Überlegungen sind darauf hinausgelaufen, ein Modell (Stumpfs Kurve als Essenz seiner Verschmelzungsversuche) mit einem Modell (der Koinzidenz rechteckige Impulse) zu simulieren. In Wirklichkeit werden von den Bewegungen der Basilarmembran Salven von Spikes ausgelöst, an denen sich eine zeitlich-statistische Verteilung erkennen lässt. Diese Salven, die dadurch entstehen, dass die Impulse nicht exakt zum gleichen Zeitpunkt getriggert werden, sind ein typisches Phänomen von Nervenbündeln; denn sie kommen nur durch das Zusammenwirken vieler vereinigter Nerven zustande. An der einzelnen Nervenfaser ist Periodizität (wegen der wiederholt auftretenden Aussetzer) kaum erkennbar. Salven von Spikes sind auch auf anderen Sinnesgebieten inzwischen Gegenstand intensiver statistischer Untersuchungen, weil sie Aufschluss über das Zusammenwirken der einzelnen Neuronen geben (Domany et al., 1994; Koch, 1999, S. 332 u. 351). Die an einer Stelle der Basilarmembran ausgelösten Salven korrelieren mit eben solchen Salven, die an anderen Stellen der Basilarmembran ausgelöst und/oder auf Umwegen mit einer – der Periodendauer der Schwingung entsprechenden – Verzögerung an eine Stelle gelangen, an der sie mit den zuvor ausgelösten Salven zusammentreffen. Diese Korrelation erfolgt selbst mit einer zeitlichen statistischen Verteilung, die aus der Breite der Spikes, der Latenzzeit und den beiden ursprünglichen Verteilungen der Salven resultiert (Fricke, 2009).

## 4 Ausblick

Nach den Registrierungen von Tramo et al. (2001) muss man annehmen: Katzen können Konsonanz von Dissonanz unterscheiden. Und man darf sicher extrapolieren: Säugetiere sind aufgrund der getriggerten Nervenimpulse in der Cochlea und deren dort schon erfolgenden Autokorrelation in der Lage, konsonante und dissonante Schallvorgänge als unterschiedliche Klangeindrücke wahrzunehmen.

Katzenohren sind während der evolutionären Entwicklung sicher nicht für musikalische Intervalle optimiert worden. Dass sie dennoch auf harmonische Verhältnisse dargebotener Töne in besonderer Weise reagieren, liegt daran, dass harmonische Intervallverhältnisse Teile des harmonischen Gerüsts der in der Natur und Tierwelt vorkommenden Klangerscheinungen darstellen. Harmonische Intervallverhältnisse sind beliebig ausgewählte Zahlenpaare aus dem unteren Teil der harmonischen Teiltöne von periodischen Schallvorgängen. Der untere Teil der Harmonischen erstreckt sich dabei bis zu der Tonhöhe, bei der mehrere Harmonische in eine Kopplungsbreite (= Frequenzgruppe) auf der Basilarmembran fallen und dort zu einem pulsierenden Bewegungsvorgang verschmelzen. Die Grenze, die für den Menschen etwa bei einer kleinen Terz anzusetzen ist, dürfte für die verschiedenen Tierarten ver-

schieden sein. Katzenohren sind in dieser Hinsicht – nicht, was den Frequenzbereich des Hörens betrifft – den Menschenohren offenbar sehr ähnlich.

In Anbetracht dieser Ergebnisse könnten viele Studien, die über den erfolgreichen und fördernden Einsatz von (harmonischer) Musik bei Mensch und Tier berichten, eine neue Bewertung erfahren.

## Literatur

- Békésy, G. v. (1943). Über die Resonanzkurve und die Abklingzeit der verschiedenen Stellen der Schneckentrennwand. *Akustische Zeitschrift* 8, 66–76.
- Békésy, G. v. (1947). The variation of phase along the basilar membrane with sinusoidal vibrations. *Journal of the Acoustical Society of America*, 19, 452–463.
- Békésy, G. v. (1962): *Paper presented at the Fourth ICA Congress*, Copenhagen 1962 (not printed in the Congress Report).
- Bergeijk, W. A. v., Pierce, J. R. & David, E. E. (1960). *Die Schallwellen und wir*. München: Desch.
- Domany, E., Hemmen, J. L. v. & Schulten, K. (Eds.). (1994). *Models of neural networks ii, temporal aspects of coding and information processing in biological systems*. New York/Berlin/Heidelberg: Springer.
- Dudel, J. (1990). Informationsvermittlung, Erregungsübertragung von Zelle zu Zelle. In R. F. Schmidt & G. Thews (Hrsg.), *Physiologie des Menschen* (S. 20–63). Berlin/Heidelberg/New York: Springer.
- Euler, L. (1739). *Tentamen novae theoriae musicae*. St. Petersburg: Academia Scientiarum.
- Eximeno, A. (1774). *Dell' Origine e delle regole della Musica*. Rom: Barbiellini.
- Fletcher, H. (1934). Loudness, pitch, and timbre of musical tones and their relation to the intensity, the frequency, and the overtone structure. *Journal of the Acoustical Society of America*, 6, 59–69.
- Fricke, J. P. (1973). Die Relativität optimaler Intervallgrößen. In C. Dahlhaus (Hrsg.), *Bericht über den musikwissenschaftlichen Kongreß Bonn 1970* (S. 397–400). Kassel: Bärenreiter.
- Fricke, J. P. (1996). Kombinationstöne. In L. Finscher (Hrsg.), *Die Musik in Geschichte und Gegenwart*, Sachteil Band 5 (Sp. 482–486). Kassel/Stuttgart: Bärenreiter.
- Fricke, J. P. (2002). Pitch bending und das Harmonium als Reininstrument. In M. Lustig (Hrsg.), *Harmonium und Handharmonika*, 20. Musikinstrumentenbau-Symposium (S. 105–116). Blankenburg: Stiftung Kloster Michaelstein (= Michaelsteiner Konferenzberichte Bd. 63).
- Fricke, J. P. (2005a). Psychoakustik des Musikhörens. In H. de la Motte-Haber & G. Rötter (Hrsg.), *Musikpsychologie* (S. 101–154). Laaber: Laaber. (= Handbuch der Systematischen Musikwissenschaft Bd. 3).
- Fricke, J. P. (2005b). Classification of perceived musical intervals. In C. Weihs & W. Gaul (Eds.), *Classification – the ubiquitous challenge, Proceedings of the 28<sup>th</sup> Annual Conference of the Gesellschaft für Klassifikation e. V., University of Dortmund 2004* (pp. 585–592). Berlin/Heidelberg/New York: Springer.
- Fricke, J. P. (2009). A statistical theory of musical consonance proved in praxis. In *Proceedings of the 32<sup>nd</sup> Annual Conference of the Gesellschaft für Klassifikation e. V.* Hamburg 2008, im Druck.
- Galilei, G. (1638). *Discorsi e dimostrazioni matematiche intorno à due nuove scienze attenenti alla mecanica ed i movimenti locali*. Leiden: Elsevier.

- Hartmann, W. M. (1998). *Signal, Sound, and Sensation*. 2. Auflage, New York: Springer.
- Helmholtz, H. v. (1863, <sup>5</sup>1896). *Die Lehre von den Tonempfindungen*. Braunschweig: Vieweg.
- Hesse, H.-P. (1972). *Die Wahrnehmung von Tonhöhe und Klangfarbe als Problem der Hörtheorie*. Köln: Arno Volk (= Veröffentlichungen des Staatlichen Instituts für Musikforschung Bd. 6).
- Hesse, H.-P. (2003). Hatte Lipps doch recht? Tonverwandtschaft und Tonverschmelzung im Lichte der heutigen Gehörphysiologie. In K. W. Niemöller & B. Gätjen (Hrsg.), *Perzeptiven und Methoden einer Systemischen Musikwissenschaft* (S. 133–144). Frankfurt a. M.: Lang. (Systemische Musikwissenschaft Bd. 6).
- Hornbostel, E. M. v. (1926). Psychologie der Gehörserscheinungen. In E. Bethe (Hrsg.), *Handbuch der normalen und pathologischen Physiologie*, Band 11 (S. 701–730). Berlin: Julius Springer.
- Jeffress, L. A. (1948). A place theory of sound localization. *Journal of Comparative and Physiological Psychology*, 41, 35–39.
- Kiesewetter, R. G. (1846). Einleitung und Zusatzartikel. In R. G. Kiesewetter (Hrsg.), *Der neuen Aristoxener zerstreute Aufsätze über das Irrige der musikalischen Arithmetik und das Eitle ihrer Temperatur-Rechnungen* (S. 1–16 u. 64–68). Leipzig: Breitkopf und Härtel.
- Koch, C. (1999). *Biophysics of Computation*. New York/Oxford: Oxford University Press.
- Kuttruf, H. (2004). *Akustik. Eine Einführung*. Stuttgart: Hirzel.
- Langner, G. & Schreiner, C. E. (1988). Periodicity Coding in the Inferior Colliculus of the Cat. II. Topographical Organization. *Journal of Neurophysiology*, 60 (6), 1823–1840.
- Langner, G. (1995). The temporal and spatial representation of periodicity pitch in the auditory system. In A. Melka (Ed.), *Proceedings of the 32<sup>nd</sup> Czech Conference on Acoustics* (pp. 17–24). Prague: Vuzort.
- Langner, G., Sams, M., Heil, P. & Schulze, H. (1997). Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. *Journal of Comparative and Physiological Psychology*, 181, 665–676.
- Langner, G. & Braun, S. (2000). Nachweis einer orthogonalen Repräsentation von Periodizitäts- und Frequenzinformation im Colliculus inferior mit der 2-Deoxyglucose Methode. In A. Sill (Hrsg.), *Fortschritte der Akustik, DAGA 2000* (S. 300–301), Bad Honnef: DPG-GmbH.
- Licklider, J. C. R. (1951). A duplex theory of pitch perception. *Experientia*, 7, 128–134.
- Licklider, J. C. R. (1959). Three auditory theories. In S. Koch (Ed.), *Psychology: A study of a science*, Study I, Vol. 1 (pp. 41–144). New York/Toronto/London: McGraw-Hill.
- Lipps, T. (1899). Tonverwandtschaft und Tonverschmelzung. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 19, 1–40.
- Louven, C. (1998). *Die Konstruktion von Musik*. Frankfurt a. M.: Lang (Systemische Musikwissenschaft Bd. 1).
- Louven, C. (2005). Reiz- und wissensgeleitete harmonische Informationsverarbeitung. In H. de la Motte & G. Rötter (Hrsg.), *Musikpsychologie. Handbuch der Systematischen Musikwissenschaft*, Band 3 (S. 208–230). Laaber: Laaber.
- Mengoli, P. (1670). *Speculationi di Musica*. Bologna: Benacci. (siehe auch: Fricke, J. P. (1973). Moderne Ansätze in Mengolis Hörtheorie. In H. Hüschen (Hrsg.), *Festschrift K. G. Fellerer zum 70. Geburtstag* (S. 117–125). Köln: Arno-Volk-Verlag).
- Mersenne, M. (1636/37). *Harmonie universelle*. Paris: Sebastien Cramoisy.
- Mertens, P. H. (1975). *Die Schumannschen Klangfarbengesetze und ihre Bedeutung für die Übertragung von Sprache und Musik*. Frankfurt/M: Bochinsky.

- Meyer-Eppler, W. (1951). Die Exhaustions-Schwingungsanalyse und ihre Beziehung zur Theorie des Hörens. *Physikalische Blätter*, 7, 355–361.
- Meyer-Eppler, W. (1959, 21969). *Grundlagen und Anwendungen der Informationstheorie*. Berlin/Göttingen/Heidelberg: Springer (= H. Wolter & W.D. Keidel (Hrsg.), Kommunikation und Kybernetik in Einzeldarstellungen, Bd. 1).
- Meyer-Eppler, W., Sendhoff, H. & Rupprath, R. (1959). Residualton und Formantton. *Gravesaner Blätter* 4 (4), 70–83.
- Miskiewicz, A. & Rogala, T. (2003). Roughness and Dissonance of Musical Dyads. In R. Kopiez, A. C. Lehmann, I. Wolther & C. Wolf (Eds.), *Abstracts of the 5th ESCOM Conference September 8–13, 2003; Hannover* (Code 108). CD-ROM. epOs-music: Osnabrück.
- Muzzulini, D. (2006). *Genealogie der Klangfarbe*. Bern: Lang.
- Ohm, G. S. (1843). Über die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen. *Annalen der Physik und Chemie*, 59, 513–565.
- Ohm, G. S. (1844). Noch ein Paar Worte über die Definition des Tones. *Annalen der Physik und Chemie*, 62, 1–18.
- Plomp, R. & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, 38, 355–361.
- Rameau, J.-P. (1722). *Traité d'harmonie reduite à ses principes naturels*. Paris: Ballard.
- Reuter, C. (2002). *Klangfarbe und Instrumentation*. Frankfurt a. M.: Lang (Systemische Musikwissenschaft Bd. 5).
- Sacerdote, G. G. (1967). Autocorrelation of periodic functions. *Acustica*, 18, 282–287.
- Schouten, J. F. (1940a). The residue, a new component in subjective sound analysis. In *Proceedings of Koninklijke Nederlandsche Akademie van Wetenschappen* 43 (3) (pp. 356–365). Amsterdam.
- Schouten, J. F. (1940b). The residue and the mechanism of hearing. In *Proceedings of Koninklijke Nederlandsche Akademie van Wetenschappen* 43 (8) (pp. 991–999). Amsterdam.
- Schouten, J. F. (1971). The Residue Revisited. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing* (pp. 41–54–58). Leiden: Sijthoff.
- Schreiner, C. E. & Langner, G. (1988). Periodicity Coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *Journal of Neurophysiology*, 60 (6), 1799–1822.
- Seebeck, A. (1841). Beobachtungen über einige Bedingungen der Entstehung von Tönen. *Annalen der Physik und Chemie*, 53, 417–436.
- Seebeck, A. (1843). Über die Sirene. *Annalen der Physik und Chemie*, 59, 449–481.
- Seebeck, A. (1844). Über die Definition des Tones. *Annalen der Physik und Chemie*, 63, 353–368.
- Shackford, C. (1962). Some aspects of perception. II. Interval sizes and tonal dynamics in performance. *Journal of Music Theory*, 6, 66–90 u. 295–303.
- Stevens, S. S. & Davis, H. (1938). *Hearing, its psychology and physiology*. New York: John Wiley and Sons.
- Stumpf, C. (1883, 1890). *Tonpsychologie*, 2 Bde. Leipzig: Hirzel.
- Terhardt, E. (1972). Zur Tonhöhenwahrnehmung von Klängen. I. Psychoakustische Grundlagen, II. Ein Funktionsschema. *Acustica*, 26, 173–186 u. 187–199.
- Terhardt, E. (1976). Ein psychoakustisch begründetes Konzept der musikalischen Konsonanz. *Acustica*, 36, 121–137.
- Tramo, M. J., Cariani, P. A., Delgutte, B. D. & Braid, L. D. (2001). Neurobiological foundations for the theory of harmony in western tonal music. In R. J. Zatorre &

- I. Peretz (Eds.), *Biological foundations of music* (pp. 92–116). New York: The New York Academy of Sciences (Annals of the New York Academy of Sciences, Vol. 930).
- Voigt, W. (1985). *Dissonanz und Klangfarbe. Instrumentationsgeschichtliche und experimentelle Untersuchungen*. Bonn: Verlag für systematische Musikwissenschaft (Orpheus-Schriftenreihe 41).
- Walliser, K. (1968). *Zusammenhänge von Hüllkurvenperiode und Tonheit bei der Bildung der Periodentonhöhe*, Diss. München: Technische Hochschule.
- Walliser, K. (1969). Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe. *Acustica*, 21, 319–329.
- Wegscheider, K. (1986). Temperierungsprobleme in der Praxis des heutigen Orgelbauers am Beispiel der großen Silbermannorgel im Freiburger Dom – einige Gedanken zu den fünf Werckmeistertemperaturen von 1691. In E. Thom (Hrsg.), *Bericht über das Werckmeister-Kolloquium aus Anlaß des 340. Geburtstages von Andreas Werckmeister am 30. November 1985* (S. 56–76). Blankenburg: Stiftung Kloster Michaelstein (Studien zur Aufführungspraxis und Interpretation von Musik des 18. Jahrhunderts, Heft 30).
- Zwicker, E. (1971). Zusammenhänge zwischen neueren Ergebnissen der Psychoakustik. In Verein Deutscher Ingenieure (Hrsg.), *Akustik und Schwingungstechnik*, Plenarvorträge und Kurzreferate der Gemeinschaftstagung Berlin 1970 (S. 9–21). Düsseldorf: VDI-Verlag.