

Was leistet unser Gehör wirklich?

Eckard Blumschein

Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik, Otto-von-Guericke-Universität Magdeburg
blumschein@et.uni-magdeburg.de

Die Erklärung des Gehörs als Frequenzanalyse des Schallsignals auf der Basilarmembran greift zu kurz und ist nicht mehr gerechtfertigt seit man beginnt, die Funktion der äußeren Haarzellen und des Gehirns zu verstehen. Will man Lehrbücher, Messungen, Vorschriften, Tontechnik, kognitive Systeme etc. mit dem Leistungsvermögen des Gehörs in Einklang bringen, kommt man an einer fundamentalen Revision jener Vorstellungen nicht vorbei, die auf ungerechtfertigten Anwendungen grundsätzlich richtiger Theorien beruhen. Einige Beispiele sollen dazu anregen, auf breiter Front das traditionelle mathematisch formalisierte Gebäude mechanisch und energetisch begründeter Vorstellungen als unzureichend zu erkennen, durch mitunter kontraintuitive neurowissenschaftlich fundierte Erkenntnisse zu ersetzen sowie daraus die wesentlichen Konsequenzen für all jene Teilgebiete der Akustik abzuleiten, die mit dem Hören zu tun haben.

1 Vergleich zwischen Gehör und Spektrogramm

Es stimmt nicht, daß wir nur deshalb aus dem Spektrogramm eines Schallsignals viel weniger ersehen können als wir hören weil uns die Übung fehlt. Ein Spektrogramm kann einerseits viel mehr Frequenzen abbilden als unser Gehör auflöst während andererseits so wichtige Merkmale wie eine virtuelle Grundwelle und Harmonierelationen in ihm überhaupt nicht in Erscheinung treten und sich meist auch nicht rekonstruieren lassen. Schon 1841 hatte Seebeck gezeigt, daß das Ohr entgegen der auf Müller 1838 zurückgehenden Auffassung Ohms kein simpler Analysator der Frequenz sein kann. Sein Einwand wurde ignoriert weil er die noch immer praktizierte Anwendung der Fourier-Analyse zur Beschreibung des Hörens in Frage stellt. Inzwischen wächst die Einsicht, daß die fehlende Grundwelle lediglich eines von vielen Indizien für die weit unterschätzte Komplexität des Hörvorgangs ist, so daß man von einer mathematischen Abbildung des Schallsignals keine adäquate Beschreibung des Hörens zu erwarten hat. Strittig sind die Fragen, in welchem Umfang Korrekturen an all dem nötig, möglich und überhaupt sinnvoll sind, was auf der alten Theorie beruht. Im Fall des Spektrogramms stören ferner besonders das Dilemma zwischen schlechter Frequenz- und schlechter Zeitauflösung wählen zu müssen und die Akausalität des üblichen Fensters der Fouriertransformation. Wavelet-Transformationen und FTT [1] bieten hierfür Abhilfe.

2 Aufklärung eines mutmaßlichen Trugschlusses

Entspricht das Spektrum den tonotopisch geordneten Signalen des Hörnervs, und ist die Basilarmembran für die Frequenzanalyse allein verantwortlich? Zweifel daran kamen auf, als man aus an Katzen gemessenen Hörnerv-Signalen nur eine grobe Frequenzanalyse ersehen konnte, während die psychoakustisch bestimmte Frequenzauflösung mit 0.2% sehr genau ist. Gravierende Differenzen zwischen traditioneller Theorie, psychoakustischen und physiologischen Befunden sind leider keine Ausnahme. Die Suche nach widerspruchsfreien Interpretationen beginnt zweckmäßig mit der Aufklärung mutmaßlicher Trugschlüsse. Es ist verständlich, daß Beweise für die bis ins Gehirn reichende Tonotopie als Stützen der traditionellen Vorstellung sehr willkommen waren. Dabei hatte man allerdings kaum Vorstellungen welche Rolle der Struktur der Nervenverbindungen im Gehirn zuzuschreiben sein könne. Ein Schlüssel zum besseren Verständnis mag die Beachtung des Umstands sein, daß Nervenzellen immer nur kurzzeitig erregt sind. Während sie "feuern" ist ihr elektrisches "Aktionspotential" hoch. Zeitliche Koinzidenzen der Aktionspotentiale können mehr oder weniger benachbarte Neuronen ebenfalls erregen oder auch hemmen. Die Signalverarbeitung im Gehirn beruht also auf der Dynamik von auch zeitlich strukturierten Mustern, wengleich die bisher realisierbaren tomografischen Abbildungen nur die zeitlich gemittelte Erregung erfassen und so ein irreführend statisches Bild suggerieren.

Deshalb ist ein bisher vernachlässigter Fakt sehr wichtig: Mit der tonotopischen Ordnung ist zugleich auch eine gemeinsame zeitliche Latenz verbunden. Die spektrale Tonalität beruht also höchstwahrscheinlich auf zeitlicher Koinzidenz [2]. Anders ist es nicht zu erklären, daß Töne gleichartig klingen, deren Frequenzen sich um Oktaven voneinander unterscheiden. Man darf also trotz Tonotopie eine zeitliche Kodierung nicht ausschließen. Die musikalische Wahrnehmung reicht bis ungefähr 4500 Hz, wo die als Phasen-Kopplung bezeichnete zeitliche Strukturierung aufhört, so daß zur Wahrnehmung höherer Frequenzen nur noch die Tonotopie beiträgt.

3 Weitere vermutete Mechanismen

Diese doppelte Kodierung in Ort und Zeit ist offenbar im Telefon-Frequenzbereich von 300 bis 3000 Hz besonders ausgeprägt und hier vermutlich nicht völlig redundant. Vielleicht vermitteln kleine Verwerfungen zwischen zeitlichem und örtlichem Kode einen Teil der Identifikation von Schallquellen und ihrer Ortung im Raum. Im Gegensatz zur arg begrenzten zeitlichen Auflösung im Spektrogramm beweist das menschliche Gehör erstaunliche zeitliche Präzision. Wer versucht zwei Töne so einzustellen, daß sie sich genau um eine Oktave unterscheiden, macht dabei stets einen gut reproduzierbaren systematischen Fehler von etwa 15 μ s. Offensichtlich sind Gruppen von Haarzellen und von Neuronen hinsichtlich ihrer zeitlichen Auflösung aufeinander abgestimmt.

Die Frequenzgruppen realisieren ein definiertes Begrenzen der spektralen Auflösung, gemäß den Erfordernissen des Erkennens von Formanten. Man darf sich also nicht täuschen lassen. Die primäre Fähigkeit des Innenohrs winzige Frequenzunterschiede aufzulösen wird durch die Breite der Frequenzgruppe verschleiert. Bekanntlich ist diese Breite so verblüffend unabhängig vom Schallpegel, daß man annahm, sie wird noch nicht vom Innenohr sondern erst vom Mittelhirn bestimmt. Etliche Anzeichen deuten darauf hin, daß die stationäre Bandbreite jeder Frequenzgruppe sowohl der Breite der Aktivität äußeren Haarzellen bei Erregung mit einer festen Frequenz entspricht als auch von der Innenohr-Mechanik bestimmt wird. Bisher galt die Basilarmembran als der frequenzanalysierende Resonator. Nachdem 1991 Dancer den Haarzellen diese Rolle zugeschrieben hatte, sah Braun 1994 die Basilarmembran nur noch als Absorber überschüssiger Energie an. Vielleicht ist sie in Wirklichkeit für die Breite der Frequenzgruppen, die sogenannte kritische Bandbreite, verantwortlich.

Man stellt sich vor, daß Sprache, Musik und andere praktisch relevante komplexe Schallsignale durch die Anordnung mehrerer gleichzeitig ausgebildeter Frequenzgruppen charakterisiert sind. Da wesentliche Einsparungen an Übertragungskapazität auf dem Weglassen maskierter Information beruhen, wünscht man sich ein besseres Verstehen der Dynamik von Frequenzgruppenbreite und Asymmetrie der Maskierung. Dies wird vor allem durch die Kompliziertheit hydromechanischer Modelle der Wanderwelle(n) behindert, die schon deshalb ungläubwürdig sind, weil man sie sehr gut an die beobachteten Phänomene angepaßt hatte, bevor klar wurde, daß letztere wesentlich von der unberücksichtigten Aktivität der äußeren Haarzellen bestimmt sind. Faßt man die Wanderwellen wieder wie von Helmholtz lediglich als das sichtbare Resultat örtlicher Resonanzen auf, die voneinander ziemlich unabhängig sind, so vereinfacht sich die Beschreibung.

4 Auswirkungen der Refraktärzeit

Eigentlich wäre zu erwarten, daß technische Sensoren den Sinnesorganen überlegen sind, denn Einzelneuronen arbeiten langsam. Der Ingenieur mag sich wundern, was die Hörschnecke so extrem schnell macht, Robustheit gegenüber Störungen vorbereitet sowie störende Abtasteffekte vermeidet. Je Bark ist in jeder Reihe Platz für etwa 170 äußere bzw. 150 innere Haarzellen, und der Hörnerv

ist gar mit mehr als 1,200 Fasern je Bark für die stochastische Kodierung ausgestattet. Das je Kanal übertragende Zeitmuster muß sich jeweils zufällig gestreut auf eine Gruppe benachbarter Neuronen verteilen, da jedes einzelne von ihnen erst nach Ablauf seiner Refraktärzeit erneut feuern könnte. Auch wenn die Streuung des spontanen Feuerns in Ruhe von der Molekularbewegung stammen mag, ist eine gleichmäßig regellose Verteilung sowie schnelle Rückkehr in diesen Zustand nach ausgeprägter Phasenkopplung keinesfalls selbstverständlich sondern setzt vermutlich neben der Varianz von Spontanfeurraten die gleichmäßige Funktionstüchtigkeit der benachbarten Haarzellen voraus. Ein scharf lokalisierter, im Audiogramm unauffälliger Funktionsausfall könnte also Tinnitus bewirken.

Intervall-Histogramme [3] weisen für Signalperioden oberhalb der Refraktärzeit gelegentliches zweifaches Feuern innerhalb einer Signalperiode aus. Die entsprechende Häufung nullter Ordnung ist bei 2,5 ms (400 Hz) gerade sichtbar und bei 5 ms (200 Hz) schon recht auffällig. Ab 8 ms (125 Hz) ist die etwa einer Bestfrequenz von 1000 Hz entsprechende Häufung nullter Ordnung höher als alle anderen, auch als das Maximum erster Ordnung. Diese Intervalle sind kürzer als die Refraktärzeit. Man hört sie also entgegen der bereits widerlegten [4] Autokorrelationshypothese nicht.

5 Zeitliche Wahrnehmung

Angenommen, die spektrale Tonalität beruht tatsächlich auf zeitlicher Koinzidenz. Dann hat man hiervon die oft als "zeitlich" bezeichnete Wahrnehmung der Periodizität zu unterscheiden, die sich ab Colliculus inferior als Periodotopie manifestiert [5]. Während beide tonotopschen Codes weite Frequenzbereiche vom Innenohr bis in den Kortex übertragen wo eine Verknüpfung zwischen ihnen erfolgen kann, sind Periodizitäten bzw. Modulationen nur hörbar, wenn ihre Periode die Refraktärzeit überschreitet. Sogar die kleine braune Fledermaus hört die Periodizität nur bis etwa 0,8 kHz [6]. Übertragen werden derart niedrige Frequenzen paradoxerweise vor allem von Nervenfasern im basalen Bereich der Cochlea mit hoher Bestfrequenz, also an der Tonotopie vorbei.

Erforschte neuronale Mechanismen sind evolutionär plausibel und zeigen wenig Ähnlichkeit mit üblichen technischen Lösungen. Eine paradoxe Vergrößerung der Latenz mit wachsendem Pegel wird anscheinend durch eine partielle Hemmung von spontanem asynchronem Feuern der Chopper-Zellen bewirkt [6]. Während Nervenzellen im Cortex ohne Signal regelmäßig feuern, ist der Zustand in dem Merkmale selektiert werden durch scheinbar chaotische Abstände zwischen Spikes gekennzeichnet. Stevens und Zador [7] bezeichnen dieses Verhalten als kontraintuitiv.

Da die Neuronen bereits im Nucleus cochlearis vielfältig spezialisiert zusammenwirken, ist die Interpretation ihrer Funktion gewagt. Grob vereinfacht mag von den drei Zweigen des Hörnervs der anteriorventrale über Büschelzellen schnelle kontralaterale Vergleiche bedienen und damit für die μ s-Auflösung verantwortlich sein, während der posteriorventrale über multipolare Chopper-Zellen nicht nur die erwähnte Periodizität aufnimmt sondern, auch als Überträger des Spektrums gilt und am dorsalen speziell onsempfindliche Neuronen hängen. Dies bedeutet, alle Zweige erfassen vorrangig zeitliche Merkmale.

Die Art wie wir die Periodizität beispielsweise einer Burstfolge hören weist Besonderheiten auf: Das Geräusch wird als fauchend bis klopfend bezeichnet. Seine empfundene Intensität korreliert mit Steilheiten im Signal. Ein Unter- oder Überschreiten der minimal hörbaren Frequenz ist nicht wahrnehmbar. Wechselt die Polarität der Steilheiten, so hört man ein dubioses Auf- oder Abklingen, je nach Vorzeichen, aber keinen stationären Unterschied. Da die Aktionspotentiale auch bei negativem Schallsignal stets positiv sind, halbiert sich die gehörte Periode einer Folge alternierender Klicks gegenüber der Grundperiode des Spektrums [8]. Auch die Phasenempfindlichkeit ist plausibel. Die neuronale Analyse der Periodizität wird anders als die schon im Innenohr erfolgende Frequenzanalyse nicht mehr von Gruppen sich zeitlich stochastisch ergänzender Zellen ausgeführt sondern primär von einzelnen, der Refraktärzeit unterliegenden Chopperzellen. Sie reagieren nicht direkt auf die Nulldurchgänge des Signals, und es trifft auch nicht zu, daß die zeitliche Feinstruktur des Signals in der "zeitlichen" Wahrnehmung erhalten bleibt.

6 Messen und bewerten wir angemessen?

Geht man davon aus, daß sich das Gehör in einem Ausleseprozeß an praktische Erfordernisse angepaßt hat, erscheint es ratsam die ausschließlich an einer angenommenen linearen Frequenzanalyse orientierte "theoretische" Begründung des Vorgehens beim Ermitteln der Leistungsfähigkeit des Gehörs grundsätzlich in Frage zu stellen. Warum berücksichtigen die bei der Prüfung von Lautsprechern geltenden Vorschriften jene Aspekte überhaupt nicht, die J. Manger als für die Wahrnehmung besonders wichtig erkannt hat? Warum verglich man [9] die Suche nach hörbarer Polarität mit der nach außerirdischem Leben? Ergibt sich aus dem zweifellos richtigen Shannon'schen Abtasttheorem und der maximal bis knapp 20 kHz reichenden Hörbarkeit des Schalls wirklich, daß Abtastfrequenzen über 40 kHz sinnlos sind? Die opponierenden Praktiker haben sich nicht geirrt. Es ist nicht gerechtfertigt, lediglich Betrag und Phase zu messen und unabhängig voneinander zu bewerten, denn beispielsweise die Korrektheit der Sprungantwort ist daraus nicht ersichtlich. Noch viel weniger ist es zulässig, die Phase stets zu ignorieren,

Skepsis ist auch geboten gegenüber der Verwendung des Begriffs akustische Energie im Zusammenhang mit der angenommenen Bewegung der Basilarmembran, denn außer bei hohem Schallpegel werden die inneren Haarzellen erwiesenermaßen durch die Aktivität der äußeren Haarzellen erregt, welche dafür Energie aus einer Spannungsquelle erhalten und das Eingangssignal nichtlinear widerspiegeln. Insbesondere wäre es unverantwortlich, die Möglichkeit der Schädigung des Gehörs nur direkt nach der Leistung des Schalls zu beurteilen. Wie andere Sinnesorgane vollbringt ja auch das Ohr eine von technischen Systemen vorläufig nicht erreichbare wesentliche Leistung damit, daß es sich ideal anpaßt und selbst schützt. Die individuelle Gefahr ist davon abhängig in welchem Maße dies funktioniert.

7 Dynamischer Dynamikbereich

Unser Gehör beweist erstaunliche Anpassungsfähigkeit. Es unterdrückt unwichtige Echos, erlaubt gezieltes Hin- oder auch Weghören und kann Eingangsleistungspegel wahrnehmen, die sich um den Faktor 1,000,000,000,000 voneinander unterscheiden, wobei sich an der Hörschwelle die Steigbügelplatte rechnerisch nur um Pikometer bewegt. Sicherlich tragen mehrere Tricks dazu bei, vor allem die spontane Aktivität der äußeren Haarzellen mit einer efferenten Gegenkopplung am Rande der Instabilität sowie stochastische Resonanz. Der leiseste Ton, den man gerade zu hören beginnt, ist rund 10 dB stärker als der gerade nicht mehr hörbare Ton. Dies deutet auf eine Resonanz hin, welche beispielsweise die akustische Impedanz lokal verringern und so die Eingangsleistung örtlich konzentrieren könnte. Wie erwartet, braucht dieses Anschwingen Zeit, allerdings nur wenige Perioden. Die evolutionäre Lösung ist also auch vorbildhaft schnell.

Literatur

- [1] E. Terhardt: Akustische Kommunikation - Grundlagen mit Hörbeispielen. Springer, Berlin/Heidelberg 1998.
- [2] E. Blumschein: Berücksichtigung der zeitlichen Wahrnehmung im Abbild natürlicher Geräusche. DAGA 2000, 308-309.
- [3] J. E. Rose, J. F. Brugge, D. J. Anderson, J. E. Hind: *Journal of Neurophysiology* **30** (1967) 776.
- [4] C. Kaernbach, L. Demany: Psychophysical evidence against the autocorrelation theory of auditory temporal processing. *J. Acoust. Soc. Am.* **104** (1998) 4, 2298-2306.
- [5] G. Langner: Periodicity coding in the auditory system. *Hearing Res.* **60** (1992) 115-142.
- [6] A. Galazyuk, A. Feng: Neural oscillation and time domain processing in the inferior colliculus. ARO Meeting 2001.
- [7] C. F. Stevens, A. M. Zador: Input synchrony and the irregular firing of cortical neurons. *Nature neurosci.* **1** (1998) 3,210-217.
- [8] R. Warren: Auditory Perception. Cambridge Univ. Press 1999.
- [9] J. Monforte, B. Karley, commenting R. A. Greiner and D. E. Melton, "Observation on the audibility of acoustic polarity," *J. Audio Eng. Soc.* **43** (1994) 3, 147-8.

Diskussionsempfehlung

<http://iesk.et.uni-magdeburg.de/~blumsche/AuditoryFunction.html>