

DISS. ETH NO. 22948

How Nonlinear Dynamics in the Cochlea Shapes Human Sound Perception

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences (Dr. sc. ETH Zürich)

presented by

Florian Gomez

Master of Science in Physics, ETH Zurich

born 10. 11. 1987

citizen of Langrickenbach, TG

accepted on the recommendation of

Prof. Dr. Ruedi Stoop, Examiner, Supervisor
Prof. Dr. Hans Jürgen Herrmann, Co-Examiner
Dr. Tobias Reichenbach, Co-Examiner

2015

Summary

In the past fifteen years, a detailed conceptual understanding of the mechanics of the mammalian cochlea has been achieved in terms of a model of forward-coupled subcritical Hopf oscillators. This ‘Hopf cochlea’ is based on the principle of small-signal amplification, exploiting the fact that systems close to a Hopf bifurcation show exactly the amplification characteristics observed in the mammalian ear. Even though the correspondence of model and experiment has been widely acknowledged, it has only been discussed to a limited extent how exactly the nonlinear characteristics of the cochlea shape the signal and what this implies for the perception of sound as a whole.

In this thesis, a thorough characterization and discussion of the nonlinear aspects of cochlear sound processing is presented. In addition to a detailed analysis of the key processes at work, there is a strong focus on the actual implications for perceptual hearing characteristics such as pitch. The work is formulated in the language of physics, but provides comparisons to biological concepts and psychoacoustical notions whenever needed and appropriate. Furthermore, most of the analysis is performed on the mesoscopic level, where one does not explicitly account for all the microscopic biochemical and biophysical details, but rather where one tries to build a model on the basis of robust generic principles, that simultaneously describe the essential aspects of the structure and dynamics of the system under investigation.

As a first step, the response of the Hopf cochlea is exhibited for different classes of signal input. Due to the lack of a superposition principle in nonlinear systems, already quite simple inputs can lead to complex excitation patterns. This is mostly a consequence of the emergence of combination tones, which are tones that are artificially generated within the cochlea. In a second step, these complex excitation patterns are studied by a characterization of the interactions of the signal components within the cochlea using a framework based on complex network theory. This not only allows for a detailed understanding of the nonlinear sound processing, but facilitates the description of the elicited complexity in the cochlea using paradigms from the study of biological neuronal networks. Specifically, the elicited responses in the cochlea are compared to avalanches in neuronal cell cultures. Based on a large number of inputs, the distribution of the size of the avalanches in the cochlea is found to follow a power-law. It is shown that this only holds when the cochlea is not biased towards certain inputs: if some of the Hopf oscillators are tuned away from bifurcation, mimicking a process of learning, the power-laws are lost. In correspondence with results from the thermodynamic formalism of dynamical systems, this suggests that power-laws characterize a ground state of general information processing systems.

In a third step, the consequences of the nonlinear interactions within the cochlea for the characterization of sounds by pitch is investigated. Through the accurate reproduction

of the well-known pitch-shift effects from psychoacoustics, it is assessed in what way the specific construction of the cochlea influences the perception of pitch. As the main result, it follows that not only the nonlinear amplification characteristics, but in a similar way also the dissipation processes in the cochlea affect pitch. Because the main contribution to dissipation is from viscous damping, the perceived pitch is found to be essentially shaped by the cochlear fluid. This highlights the importance of our Hopf-cochlea among the mesoscopic cochlea models. Together with the demonstration of the faithful transmission of the sound from the cochlea to the cortex via the auditory nerve, the results underline the fact that pitch is already fully present at the level of the cochlea and not, as is often assumed, generated higher up in the neural system.

Taking pitch as the guiding feature, it is then finally investigated to which extent the cochlea can be tuned towards signals of interest. Using a biologically motivated protocol to modify the bifurcation parameters of the Hopf cochlea, it can be seen that despite the complex nature of the cochlea, a simple tuning is possible and feasible. To a significant degree, this result is based on the fact that tuning not only allows the suppression of unwanted frequency components, but at the same time removes associated cross-source combination tones. This strongly suggests that cochlear tuning acts as a first step of active listening. Taking this feature into account is thus of utmost practical importance for the restoration of hearing in humans.

Zusammenfassung

In den vergangenen fünfzehn Jahren ist mithilfe eines Modells von vorwärts-gekoppelten subkritischen Hopf-Oszillatoren ein detailliertes konzeptuelles Verständnis der Mechanik der Cochlea von Säugetieren erreicht worden. Dieses „Hopf-Cochlea“-Modell basiert auf dem bekannten Prinzip der Verstärkung von schwachen Signalen („small-signal amplification“) und macht damit von der Tatsache Gebrauch, dass Systeme in der Nähe einer Hopf-Bifurkation genau diejenigen Eigenschaften aufweisen, welche für die Verstärkung von Tönen in der Säugetier-Cochlea charakteristisch sind. Obwohl die Übereinstimmung von Modell und Experiment allgemein anerkannt ist, wurde bisher nur in beschränktem Masse diskutiert, in welcher Art die nichtlinearen Charakteristiken der Cochlea das Tonsignal verändern und welche Auswirkungen dies für die allgemeine Wahrnehmung von Tönen mit sich bringt.

In dieser Dissertation wird eine gründliche Charakterisierung und Diskussion der nicht-linearen Aspekte der Tonverarbeitung in der Cochlea vorgestellt. Neben einer detaillierten Analyse der wichtigsten Prozesse liegt der Fokus dabei auf den effektiven Auswirkungen für die charakteristischen Eigenschaften der Tonwahrnehmung wie der Tonhöhe. Die Arbeit ist in der Sprache der Physik gehalten, verweist aber, wann immer es nötig und passend erscheint, auf biologische Konzepte und psychoakustische Begriffe. Die Analyse geschieht zudem auf der mesoskopischen Ebene, wo man nicht auf alle biochemischen und biophysikalischen Details explizit eingeht, sondern versucht, ein Modell auf robusten und allgemein gültigen Prinzipien aufzubauen, welche ebenso die zentralen Aspekte der Form und Dynamik des untersuchten Systems beschreiben.

In einem ersten Schritt werden die grundlegenden Antwort-Eigenschaften der Hopf-Cochlea bei Stimulation mit verschiedenen Klassen von Inputs dargestellt. Da bei nichtlinearen Systemen das Prinzip der Superposition nicht gilt, hat man es schon bei einfachen Stimulationen oft mit komplexen cochleären Anregungsmustern zu tun. Dies ist mehrheitlich eine Folge der Entstehung von Kombinationstönen. Dies sind Töne, die im Ohr künstlich generiert werden. In einem zweiten Schritt werden die komplexen Anregungsmuster mittels einer Charakterisierung der Interaktionen der Signalkomponenten analysiert, wobei Begriffe aus der Theorie der komplexen Netzwerke verwendet werden. Dies ermöglicht nicht nur ein detailliertes Verständnis der nichtlinearen Tonverarbeitung, sondern auch eine Beschreibung der generierten Komplexität in der Cochlea mithilfe von gebräuchlichen Paradigmen aus dem Feld der biologischen neuronalen Netze, wobei insbesondere die cochleären Anregungen mit den bekannten „Lawinen“ in neuronalen Zellkulturen verglichen werden. Wird eine grosse Zahl an Inputs verwendet, so findet man, dass die Grösse der cochleären Anregungen einem Potenzgesetz gehorcht. Hier wird gezeigt, dass dies aber nur gilt, wenn die Cochlea keine bestimmten Inputs bevorzugt: Werden im Rahmen eines modellierten Lernprozesses einige der Hopf-Oszillatoren von der Bifurkation weggebracht, so verschwindet das Potenzgesetz. In Übereinstimmung

mit Resultaten aus dem thermodynamischen Formalismus der dynamischen Systeme lässt sich daraus schliessen, dass Potenzgesetze einen Grundzustand der Informationsverarbeitung charakterisieren.

In einem dritten Schritt werden die Konsequenzen der nichtlinearen Interaktionen in der Cochlea für die Wahrnehmung der Tonhöhe untersucht. Durch die akkurate Reproduktion der bekannten psychoakustischen Pitch-Shift-Effekte zeigt sich, inwiefern die spezifische Konstruktion der Cochlea die Tonhöhenwahrnehmung beeinflusst. Als Hauptresultat lässt sich folgern, dass es nicht nur die nichtlinearen Verstärkungseigenschaften sind, welche auf die Tonhöhe einwirken, sondern im selben Masse auch die cochleären Dissipationsprozesse. Da der grösste Teil der Dissipation durch die viskose Dämpfung zustande kommt, folgt daraus, dass die Tonhöhenwahrnehmung stark von der Präsenz der cochleären Flüssigkeit geprägt ist. Diese Feststellung streicht die spezielle Bedeutung unserer Hopf-Cochlea unter den mesoskopischen Cochleamodellen heraus. Zusammen mit der hier ebenfalls gezeigten Signaltreue der Übermittlung des cochleären Signals zum Cortex mittels dem auditorischen Nerv führt dies dazu, dass die Tonhöhe bereits vollumfänglich in der Cochlea vorhanden ist, im Gegensatz zu der verbreiteten Annahme, dass sie erst im Zentralnervensystem entsteht.

Durch den Gebrauch der Tonhöhe als Referenzeigenschaft wird im letzten Teil dieser Dissertation untersucht, in welchem Ausmass die Cochlea „getunt“ oder „eingestimmt“ werden kann, so, dass bestimmte Zielsignale verstärkt wahrgenommen werden können. Dabei werden die Bifurkationsparameter der Hopf-Cochlea in einer biologisch motivierten Weise verändert, wobei sich zeigt, dass solch ein Tuning trotz der komplexen Natur der Cochlea durchaus einfach und realistisch ist. Zu einem grossen Teil kommt dieses Resultat dadurch zustande, dass das Tuning nicht nur die unerwünschten Frequenzkomponenten unterdrückt, sondern gleichzeitig auch gerade die dazugehörigen Misch-Kombinationstöne entfernt. Dies bekräftigt die Vermutung, dass das cochleäre Tuning einen ersten Schritt des aktiven Zuhörens darstellt. Für die Wiederherstellung der Hörfähigkeit beim Menschen ist die Berücksichtigung dieser Eigenschaft daher von grösster praktischer Bedeutung.