Chapter 1
Introduction

As a branch of science matures, it often develops a synthetic component. Although many open issues remain, advances in auditory physiology, psychophysics, and theoretical modeling offer detailed information on the structure, function, and fundamental limitations of the early auditory system. Creating an artificial early auditory system, in a physical medium that shares many of the strengths and weaknesses of the biological substrate, requires a concrete specification of auditory neural computations, and provides a challenge to the field different from those of computer simulation and theoretical modeling. My co-workers and I are building analog VLSI (very large-scale integrated) silicon circuits that model the early auditory system, both to explore the general computational principles of auditory neuroscience, and to create potentially useful devices for sound understanding, for sound localization, and for prostheses for the deaf. This dissertation documents some of our research.

1.1 Motivation for the Scientist: Why Use Analog VLSI Technology?

We believe that any medium imposes a direction on modeling; it is thus advantageous to choose a modeling medium that shares characteristics with the system under study. Analog VLSI and neural systems are different in detail, but the frameworks for computation in the two technologies are remarkably similar. Both media pack a large number of imperfect computational elements into a small space. Systems in both technologies must confront these imperfections, not as a second-order effect, but as a prerequisite for a working design. Both media offer a rich palette of primitives in which to build a structure; nonlinearities are resources for improved system performance. Neural and silicon systems are both ultimately limited not by the density of devices, but rather by the density of interconnect. Modeling neural systems directly in a physical medium subjects the researcher to many of the same pressures faced by the nervous system over the course of evolutionary time.

The design style of the projects described in this dissertation accentuates the similarities between analog VLSI and neural systems. The circuits use the full analog nature of transistors, exploiting the exponential functions inherent in the devices. The chips use voltage, current, and charge as analog data representations of system quantities; the systems also use a pulse representation that models biological action potentials. All communication in the systems is asynchronous; there are no clock signals. The speed of processing is matched to the time scale of auditory signals. Each subcircuit is dedicated to modeling a known or proposed structure in the early auditory system; circuits are not time-multiplexed between multiple functions.

Analog VLSI systems, using this design style, have another distinct advantage: real-time performance. All the systems described in this dissertation work in real time. Real-time performance allows large, complex models to be tested with a large library of input stimuli. Real-time performance also encourages the construction of electromechanical systems to test sensorimotor models.

Analog VLSI technology is an excellent medium for auditory modeling. Processing time-varying information fully exploits the real-time performance of analog VLSI systems; modeling human audition requires the processing of time-varying waveforms at frequencies up to 20 kHz. In addition, many representations of early audition do not require wiring in three dimensions, and map well to a two-dimensional silicon die.

1.2 Motivation for the Engineer: Why Model Audition?

Why should an engineer study and model early biological auditory processing, and incorporate these models into practical sonar and speech-recognition systems? The early representations of a sensory system are crucial. Early representations serve to reduce the amount of sensory data, while preserving information essential for the task of the system. Moreover, early representations express this information in a way that simplifies the task of later modules in the system. Because early biological auditory representations are the first stage of the best speech-recognition system known, it makes sense to incorporate these representations into human-made speech recognizers. Likewise, these representations are the sonic substrate for animals specialized for passive and active sonar. The sonar performance of some of these animals exceeds the performance of human-made sonar systems; incorporating biologically inspired early auditory representations into human-made sonar systems may decrease this performance gap.

Practical speech-recognition and sonar systems often require real-time performance. Modeling early biological auditory representations using conventional digital computers, in real time, is not possible with existing technology. Analog VLSI technology is a medium that offers real-time computation of these representations. Specialized digital architectures are another way to compute these representations in real time. It is beyond the scope of this dissertation to judge which method offers the best engineering solution to the problem. As detailed in the Section 1.1, analog VLSI technology, by its similarity to the biological substrate, aids the scientist in understanding existing representations and inventing new models to explain neural data. In the same way, analog VLSI technology may aid the creative engineer in designing innovative solutions to engineering problems.

1.3 Overview of the Dissertation

This dissertation describes silicon models of three aspects of early audition. Chapter 2 describes an integrated circuit that models the evoked responses of the auditory nerve. The auditory nerve is the first neural representation of audition; its carries information from the cochlea, the sense organ of hearing, to the brain. Comparisons of chip and physiological responses to classical stimuli show a qualitative similarity, with quantitative differences.

Chapter 4 describes an integrated circuit that computes the spatial position of a sound in the azimuthal plane. The chip models a neural representation of auditory space found in the barn owl, an animal specialized for passive auditory localization. The 220,000-transistor chip models the processing of two cochleas and several brainstem nuclei, computing an output map of azimuthal space by processing interaural time differences.

Chapter 5 describes an integrated circuit that models human pitch perception. Chip output approximates human performance on a variety of classical pitch-perception stimuli. Some of the circuits on the chip correspond to known physiological computations; other parts of the chip model proposed neural structures.

Inhibition is an essential element of the auditory-localization and pitch-perception chips; Chapter 3 describes the nonlinear-inhibition circuit used in both chips. This novel circuit needs only two transistors for each neuron, and requires a length of interconnect that increases linearly with the number of neurons. Chapter 6 offers suggestions for future research.