1999

A Brain for Language and the Language of the Brain

Guillermo A. Cecchi

A brain for language and the language of the brain

A thesis presented to the Faculty of
Rockefeller University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

by

Guillermo A. Cecchi

Para Silvia, que sobrellevó infinitas graduaciones, y que guarda el tesoro más precioso,

para Mamá y Papá, cuyos sacrificios, aún después de tantos años, no dejan de asombrarme,

a mi hermano, mis nuevos padres y hermanos, a todos los hermanos del alma allá en el pago, a todos los compañeros, y a mis héroes privados, Robi, Ernesto y Dani.

Abstract

This thesis consists of three sections. In the first one, we overview the basic concepts and results of computation theory, the relationship with formal language theory, and present an introduction to the possible applications to different biological systems. The second section deals with the role of noise in the transmission of information in neurons. We show that a key component of the dynamics of action potential generation in neurons, the rectification of ionic flow and negative incremental resistance to ionic gradient, can occur in Brownian transport in the environment of the cell by a purely entropic effect, i.e. driven by thermal noise. We also focus on the problem of temporal coding of information by neurons. We show that a wide class of models of spiking elements display a high sensitivity of the output message noise to the particulars of the input message. We show that predictions based on these models are corroborated by measurements of spiking activity in the visual geniculate of cats. The last section is centered in the problem of coding of complex spaces by the brain. We investigate a case of animal communication, that of songbirds. We show that there exists in the auditory brain of canaries a topographic representation of song, whereby song components are mapped into unique patterns of activity, preserving stimulus identity as well as inter-stimuli relationships. We postulate that this representation reveals a song code embedded in the topography of neural space. We discuss different theories of neural code, their relevance to the findings of syllabic represen-

tation, and preliminary approaches to models of brain function compatible with these results. These approaches are based on the hypothesis of the existence in auditory areas of birdsongs of a self-organizing neural network that, driven by statistics of the input, constructs a consistent representation of the auditory input ensemble in neural space

Contents

Acknowledgements

Science is a social phenomenon, a tangled web of people an ideas, and indeed it could not be otherwise. The work presented in this thesis is itself an acknowledgement to all the persons that directly or indirectly participated in it. However, from an individual point of view, I have to acknowledge a number of people that have contributed to my intelectual and personal growth: my advisor Marcelo Magnasco for offering me the opportunity to do science with joy, for his intelectual honesty, and for his uncompromising friendship; Sidarta Ribeiro and Claudio Mello for sharing their work and ideas without restrictions; their disposition allowed us to complete a work we all are proud of; Albert Libchaber for his permanent concern and support, and together with Mitchell Feigenbaum, for creating a space of unlimited academic freedom; to Jim Hudspeth for his invaluable support in difficult times; to Dante Chialvo for his contagious passion, his epic ideas and his friendship; to Roy Crist and again Sidarta for the long nights of ludicrous creativity, for sharing the dream and the burden of the scientist as a responsible intelectual, and for their friendship; Tim Gardner and Mariano Sigman for partaking of and

rejoicing in the whirl; Didier Chatenay for his advice and golden heart; finally, all the friends collected these years, marvelous, extravagant, quixotic, unique, denoted by their *nom de guerre:* la Armada Brancaleone.

Prologue

In the fall of 1917, I entered Haveford College with two strings to my bow - facility in Latin and a sure foundation in mathematics. I "honored" in the latter and was seduced by it. That winter Rufus Jones called me in. "Warren," said he, "what is thee going to be?" And I said, "I don't know." "And what is thee going to do?" And again I said, "I have no idea; but there is one question I would like to answer: What is a number, that a man may know it, and a man, that he may know a number?" He smiled and said, "Friend, thee will be busy as long as thee lives." I have been, and that is what we are here about.

W. S. McCulloch

The question whether objective truth can be attributed to human thinking is not a question of theory but is a practical question. Man must prove the truth – i.e. the reality and power, the this-sidedness of his thinking in practice. The dispute over the reality or non-reality of thinking that is isolated from practice is a purely scholastic question.

K. Marx, 2nd Thesis on Feurbach

There are two major sources of inspiration for the work contained in this thesis. One is the riveting development of mathematical logic and philosophy of language initiated during the last decades of the 19th century with the work of Frege and Cantor, culminating with the formalization of the concept of computation by Gödel and Turing in the 1930's. The consequences of this

conceptual revolution are still unfolding. Computers are bringing changes to many aspects of life: the way we work, the way we communicate, and even the way we think. The impact of the computer age in science is enormous, from laboratory work to computational theories of cell function and computational theories of the mind. Another fascinating aspect of language is that it is both the object of study of different disciplines like linguistics (this field in particular has been highly influenced by computer theory after the Chomskian "revolution" of the 1950's), neurology, ethology, and at the same time the scaffold on which different analytical tools like logics and computer theory have developed.

The other source of inspiration is the equally vigorous expansion of biology during the second half of the 20th century, most prominently after Watson and Crick solved the double helix structure of DNA. This particular fact marks a divide in the history of biology; questions like the origin of life on the one side, and the nature of thought on the other, which have occupied the best minds for centuries, seem now to be approachable or thinkable within a materialistic paradigm. Various molecular techniques as well as global imaging procedures have dramatically changed the way experimental biology is practiced. However, no similar development has taken place in the theoretical understanding of biological phenomena, with the exception of versions of Darwinism adapted to the current knowledge of molecular and cell biology. Biology, as seen from the more formal disciplines like physics

and chemistry, preserves a descriptive rather than explanatory quality. The complexity and variety of biological phenomena defy any formal description; life always seems to escape definitions and restraints.

These inspirational streams converge in two questions that have been the prime movers of the theoretical and experimental work presented here: *can we use language as a model of the symbolic manipulation of information in biology?*, and *what are the solutions implemented by biological systems to achieve this symbolic manipulation?* We hope that by modeling living systems as computational and informational devices, we will bridge the gap between their infinitely complex descriptions and the scattered theoretical explanations. However, we have to solve several problems along the way, amongst which are the constraints imposed by the biological substrate.

This thesis consists of three main Parts, and two bonus Addenda. In Part I (Chapter 1) we will overview the basic concepts and results of computation theory, the relationship with formal language theory, and an introduction to the possible applications to different biological systems.

Part II deals with the role of noise in the transmission of information in neurons. In Chapter 2 we will show that a key component of the dynamics of action potential generation in neurons, the rectification of ionic flow and negative incremental resistance to ionic gradient, can occur in Brownian transport in the environment of the cell by a purely entropic effect, i.e. driven by thermal noise. We will focus in Chapter 3 on the problem of temporal

coding of information by neurons. We will show that a wide class of models of spiking elements display a high sensitivity of the output message noise to the particulars of the input message. We will show that predictions based on these models are corroborated by measurements of spiking activity in the visual geniculate of cats.

Part III is centered in the problem of coding of complex spaces by the brain. In Chapter 4 we will investigate a case of animal communication, that of songbirds. We will show that there exists in the auditory brain of canaries a topographic representation of song, whereby song components are mapped into unique patterns of activity, preserving stimulus identity as well as inter-stimuli relationships. We will postulate that this representation reveals a song code embedded in the topography of neural space. In Chapter 5, we will discuss different theories of neural code, their relevance to the findings of Chapter 4, and preliminary approaches to models of brain function compatible with these results.

Finally, for the interested reader, we provide two bonus chapters with a discussion of experimental and theoretical issues of computation with biomolecules.

Part I

Formal and natural languages

Chapter 1

Algorithms, formal languages and machines

1.1 Introduction

The concept of algorithm is basically an intuitive one, as is the case with many other concepts in the very basis of the scientific reasoning. Intuitively speaking then, an algorithm is a procedure to secure the solution of a problem. Examples of problems for which well-known algorithms exist are: the number resulting from the addition of any two natural numbers, the number resulting from the multiplication of any two real numbers, the one-digit integer in the *n-th* position of the decimal expansion of π. For each one of these examples, there are several different implementations of algorithms that solve the problem, and it is easy to show their equivalence. From these simple examples, it is also clear that "problem" can be replaced by the more formal notion of *function*, and that "solving a problem" is equivalent to *computing a function*, so that the addition and the multiplication problems become

two-variable functions, and the π-expansion problem becomes a one-variable function. What happens with more complex functions, like *are there any occurrence of five 1's in a row in the decimal expansion of* π? Is it always possible to find an algorithm to compute this function, or *any* function for that matter? Is it possible to define what an algorithm should be, or equivalently, what algorithms have in common? In other words, is it possible to formalize the notion of algorithm? [1]

In trying to answer these questions, logicians and mathematicians have developed what is nowadays known as *computability theory*, which we will briefly describe in the next sections. We will discuss the idea of *model of computation* as a set of typified procedures to implement an algorithm. We will focus on the most studied model of computation, the Turing Machine, although there are many others that may be suitable for different situations (among others, Markov algorithms, Post systems, Grammars, L-systems); and we will also show the close relationship between algorithms and formal languages. Of the many results obtained in computability theory, we will highlight two: the class of functions computable by different models is unique, and by the same token, there are functions that cannot be computed by any model, forming the class of *unsolvable problems*. We will also introduce the notion of *computational complexity* as a measure of the intrinsic difficulty of different solvable problems, and *algorithmic complexity* as a mea-

[1] For further reading, see [Moll 88, Rogers 67, Minsky 67, Salomaa 73].

sure of the difficulty in the description of an object. Finally, we will discuss the generative approach to the formalization of natural languages, and the relevance of extending these ideas to the realm of biology in general.

1.2 Computable Functions

The intuitive notion of algorithm leads immediately to the concept of recursive function: the values of the function for given arguments are related to values of the same function for "simpler" arguments [2]. The simplest class of functions described in terms of recursion is that of the primitive recursive functions. This class is defined as the intersection of the classes determined by:

(i) Constant functions,

$$f(x_1, X_2, ...X_k) = m$$

(ii) Successor functions,

$$f(x) = x + 1$$

(iii) Identity functions,

$$f(x_1, x_2, ...x_k) = x_i$$

[2] A typical example is the Fibonacci sequence:

$$f(0) = 1$$

$$f(1) = 1$$

$$f(n) = f(n-1) + f(n-2)$$

(iv) Given the functions f and g_i, already in the class, the function

$$h(x_i,x_2,...x_m) = f(g_1(x_l,...x_m), g_2(...), ..., g_k(...))$$

(v) Given h and g already in the class, the function defined by

$$f(0, x_2, ...x_k) = g(x_2, ...x_k)$$

$$f(y+1, x_2, ...x_k) = h(y, f(y, x_2, , , , x_k), x_2, ...x_k)$$

It is possible to show that most of the functions customarily used in ordinary mathematics are primitive recursive, though this class is not inclusive enough to account for some functions defined by means of algorithms. Furthermore, there is a systematic method to produce an algorithmic function falling outside any given formally characterized class of algorithmic functions: it is precisely the diagonalization method. We can easily set up a list of the algorithms in a model, i.e., arranging them in the increasing size of the derivation. Let us denominate Q_x the algorithm whose position in the list is x. The corresponding function will be $g_x(x)$. We can define h such that

$$h(x) = g_x(x) + 1$$

which has an obvious algorithm. In consequence, there should exist somewhere in the list the corresponding formal algorithm for h, and let us call it g_{x_0}. Then, we would have

$$g_{x_0}(x_0) = h(x_0) = g_{x_0}(x_0) + 1$$

a contradiction which arose because we assumed h as a member of the list. There is a way to avoid the diagonalization problem: we can extend the class of functions to include *partial functions*, i.e., functions whose domain is a subset of R^k. In this way, the diagonal argument does not lead to a contradiction because $h(x)$ has not necessarily to give an output for every input.

A very important concept related to computable functions is a thesis proposed by the logician Alonzo Church, after the work of Gödel and Turing formalized the intuition of computability: known as *Church's thesis*, it postulates that the class of primitive recursive functions includes all the functions than can be computed, that is to say, it formalizes the concept of computability. This thesis has been much debated, inside and outside mathematics, because one of its consequences, or at least one of its interpretations, is that it mechanizes mathematical thinking: if something is computable, then it is partial recursive, and therefore there is a 'mechanical' algorithm to compute it, without reference to informal properties inherent to the mind, as opposed to the machine. We will return to Church's thesis later in the chapter.

1.3 Turing machines

A model of computation is an attempt to formalize both the notion of algorithm and of partial function computable by an algorithm. The first such at-

tempt was made by Turing (though, as we mentioned before, there are many of them). The Turing Machine (TM) consists of a tape which is marked off into cells of equal size. A device is able to read the symbols printed on the tape $\mathcal{S} = B, S_1, S_2, ...$, where B is the blank symbol. The device has internal states $\mathcal{Q} = q_1, q_2, ...$, and the computation proceeds according to the following rule: the device is in the internal state q_i, it reads the tape symbol S_j, and determined by the pair (q_i, S_j), it can:

(i) replace (print) S_j by S_k, go into state q_l, move one step to the right,

(ii) replace S_j by S_k, go into state q_l, move one step to the left,

(iii) halt, if q_l is the special internal state *stop*.

At any moment in time, there is only a finite number of non-blank symbols, the B's fill the rest of the cells. The initial distribution of non-blank symbols corresponds to the input to the TM. When the TM stops, the output is read in the tape; eventually, the computational problem can be whether the TM stops or not. One can think of a very simple set of symbols $\mathcal{S} = B, 1, P$, so that partial function of k-variables is associated to each TM by representing the input as 1's followed by P's separating the variables. In this way, the inputs with no outputs (i.e. non-halting) correspond to the arguments that give no value in the partial function. Each stage in a TM computation can be described by giving (i) the condition of the tape, (ii) the internal state, and (iii) the cell being examined. This determines the *instantaneous description* for the calculation, expressed as a string consisting

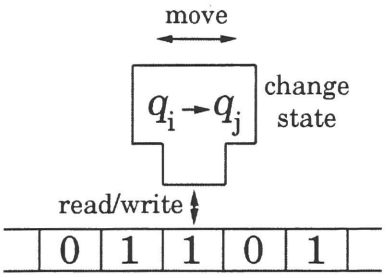

Figure 1.1: Turing Machine

of the symbols on the tape and the internal state in the position being read:

$$BBB1111P1111P111...q_i...11P1..BBB$$

Therefore, the course of a computation can be described as a finite sequence of strings, if it gives an output, or an infinite one otherwise.

Turing's formulation of algorithms as computers led to the foundation of computer science around the problem of *computability theory*, which tries to answer the question of what is computable and what is not, and what is the computational power of different computers. The basic results are the following:

1) The classes of computable functions determined by different models belong to the same class, that of the partial recursive functions. In the Turing model, it is possible to increase the number of tapes and reading devices, to allow non-deterministic transitions (when the next step in the computation is multi-valued), and yet the class remains the same.

2) This class seems to be inclusive enough: all the known intuitively

algorithmic functions have been proved to be partial recursive functions.

3) The translation of a set of instructions in a model into a set of another model is algorithmic, i.e., there is an effective procedure to translate between models.

4) Universality: let $\phi_x(y)$ be the algorithm whose position in the list (in some model) is labeled by x. As x increases, the complexity (understood as its length) of the program increases too. The universality theorem states that for any model there exists a z such that for every x and y

$$\phi_x(y) = \phi_z(x, y)$$

In other words, there exists a critical degree of complexity beyond which any further complexity can be mapped into increasing size of the input, and therefore, of the number of steps required to compute. In a Turing context, ϕ_z is called the *universal Turing machine*. In every-day computation, it is called *compiler*: no matter how complex is the program being computed, the compiler translates it into a fixed number of basic operations. This is perhaps one of the strongest and most surprising results in computer theory [Turing 36].

5) Coding: Related to universality, the Kleene's s-m-n theorem shows that for all positive integers m and $n \geq 1$, there exists a total recursive function s_n^m of $m + 1$ variables, such that for all $x, y_1, ..., y_m$:

$$\phi_x(y_1, y_2, ..y_m, z_1, ...z_n) = \phi_{s_n^m(x,y_1,...,y_m)}(z_1, ...z_n)$$

From which it is straightforward to derive the theorem of composition: there is a function g of two variables such that for all x, y,

$$\phi_{g(x,y)} = \phi_x \phi_y$$

where $\phi_x \phi_y$ stands for $\phi_x(\phi_y(z))$. This theorem is used to prove that the addition, composition, multiplication, etc., of recursive functions is recursive.

6) The halting problem: it is clear from the definition of partial recursive function that for some inputs, the computation never ends, unless the function being computed is total (infective). It turns out that there is no recursive function to decide whether a given input for a given function will halt the computation in finite time. That is to say, there is no recursive algorithm to solve the halting problem, and assuming Church's thesis, there is no algorithm at all to do it. It is worth to mention the relation of the halting problem to one of the theory of algorithmic complexity. While typical undecidable (non- algorithmic) problems like halting would imply infinite computation time, the degree of complexity of a decidable problem is measured according to the time required to solve it. Let us state the halting problem in this way: given a Turing machine M_n with semantics (the function that encodes) ϕ_n and input x, is $\phi_n(x)$ defined? In other words, does M_n ever halt if started in the initial state q_0 scanning x ? To have an algorithm for this problem would be very helpful, because it would save a lot of effort by telling us beforehand that the machine will never stop. The first attempt to solve this is using the Universal Turing Machine, with input (M_n, x). When the UTM halts, a 1 is

printed to signify that $M_n(x)$ halts. So the UTM should print a 0 to signify that $M_n(x)$ does never halt, but this implies that the UTM halts, and in such a way is not universal because is not giving the same output as M_n. Put in another way: how long are we going to wait until printing 0? This approach does not work, but clearly sticks the arrow of time in the heart of logic! A more formal approach is more involved. First, we already proved that: *there is no total Turing-computable function g which enumerates the total Turing-computable functions in the following sense: the Turing-computable function ψ is total if and only if ψ equals the function $\phi_{g(n)}$ computed by $M_{g(n)}$ for some n.* Keeping this in mind, we can state the theorem:

Consider the halt function, halt(x): $\mathbf{N} \to \mathbf{N}$ *defined as:*

$$halt(x) = \begin{cases} 1 & \text{if } M_x \text{ halts on input x} \\ 0 & \text{if } M_x \text{ never halts} \end{cases}$$

Then this is not Turing-computable. It must be noted that $halt(x)$ is a total function, and the theorem says that there is no algorithm to compute it. We can demonstrate it as follows: We can construct an $\mathbf{N} \times \mathbf{N}$ array, setting the (i, j) entry to 1 if $\phi_i(j)$ halts, and 0 if not. Using yet again Cantor's argument, we can find a function that is not in the set ϕ_i:

$$\psi(x) = \begin{cases} 1 & \text{if } \phi_x(x) = 0 \\ 0 & \text{if } \phi_x(x) = 1 \end{cases}$$

This means that we are "flipping" the values of ψ along the diagonal; but it also means that

$$\psi(x) = \begin{cases} 1 & \text{if } \phi_x(x) = 0 \\ 0 & \text{if } \phi_x(x) = 1 \end{cases}$$

By means of which, if *halt* is computable, then ϕ is too. However, if that were the case, $\psi = \phi_k$ for some k in the array, but if $\phi_k(k) = 1$ then $halt(k) = 1$ and $\psi(k) = \phi_k(k) = 1$; and if $\phi_k(k) = 0$ then $halt(k) = 0$ and $\psi(k) = \phi_k(k) = 1$, a contradiction that arose because we assumed *halt* was computable.

This result was first attained by Gödel in a more formal context. He showed that any language powerful enough to express theorems of Number Theory will be *incomplete*: there are sentences belonging to the language that cannot be deduced, yet they are valid , i.e. true. That is to say, the problem of deciding whether any sentence can be deduced in a language is equivalent to the halting problem. This is *Gödel's first incompleteness theorem*, where he also first introduced the idea of transforming the sentences of a language into natural numbers (known as Gödel's numbering), which in the context of Turing Machines is equivalent to blurring the distinction between program and data.

Now we can revisit Church's thesis: it claims that the intuitive concepts of algorithm and function computable by algorithm are fully formalized by model of computation and partial recursive function, respectively. In particular, the thesis postulates that the halting problem cannot be solved by any computer, no matter how powerful, and therefore it cannot be computed at all. Although the halting problem seems cumbersome and of interest only for computer scientists or logicians, it has had a strong impact in different areas of science, on the one hand establishing the limits of computability,

and on the other hand sparking endless debates over its consequences on mechanical theories of the mind. Regarding the former, any problem that can be mapped to the halting problem is proved unsolvable, including the group isomorphism problem (are any two arbitrary groups isomorphic?) and the tiling problem (given a finite arbitrary collection of polygons, is it possible to tessellate the plane with it?). As for the latter, detractors of the theory of the mind as a computer have used computational unsolvability as an argument to demonstrate that the mind is superior to the computer. According to this idea, Ramanujan's episodes of mathematical inspiration are evidence of non-algorithmic computation, something that would be impossible for mechanical devices.

1.4 Language and Machines

The aim of this section is to show the relation between formal languages and machines. A formal language is a system consisting of an alphabet, words over that alphabet, an alphabet of auxiliary letters, a set of rewriting rules or grammar, and a start letter. Under the action of the rules, the grammar produces some strings of symbols; those which do not have auxiliary letters qualify as the words of the language. An example: let the alphabet be a, b the start and auxiliary letter S, and the productions $S \to aSb$ and $S \to ba$. A typical derivation is:

$$S \to aSb \to aaSbb \to aababb$$

A language is recursively enumerable if there is an effective procedure, i.e., an algorithm, to enumerate the words that belong to it. It has been proved that any recursively enumerable language can be generated with the following minimal set of rules:

expansive productions

$$A \rightarrow BC$$

context-sensitive productions [3]

$$AB \rightarrow CD$$

erasing (contracting) productions [4]

$$A \rightarrow \lambda$$

terminating productions

$$A \rightarrow a$$

There is a deep relation between recursively enumerable languages and partial recursive functions, from a set-theoretical point of view. PRF are not only defined in a recursive way, but can be enumerated in the same way. This is also the case for RE languages: their elements are recursively generated, and there is a way to enumerate the languages themselves. At this point, it

[3]Context-free or sensitive does not mean "context" in any philosophical way; it means that a given letter can be rewritten in many different ways, according to neighboring letters.

[4]λ is the null symbol or zero of strings; it means that the symbol is eliminated and not replaced.

is adequate to say a few words about the auxiliary words. The use of these symbols is originated in the work by Chomsky on formalization of natural languages. They are intended to represent "nouns", "articles", "verbs", that ultimately have to be translated into actual words (we will return to this). From a more mathematical point of view, a mechanism to terminate a production is required, although there might exist non-terminating derivations. L-systems are a class of algorithms inspired in biology. The goal of these systems is to display parallel rewriting: in each step, all the letters in the string must be rewritten. The relevance to biology is straightforward, as L-systems are intended to be models of growth patterns. Moreover, context-free L-systems are suitable for the earlier stages in development, in which the cells have weak communication with each other. These systems consist of an alphabet, a start word, auxiliary letters, and a finite set of rules of transformation of the letters. EIL-systems are those which include productions of the form:

$$(\beta_1, a, \beta_2) \rightarrow \alpha$$

where the final transcription of the letter a into the word α depends upon the neighboring words β_1 and β_2 (context-sensitive). For these systems it has been proved the following theorem: a language is recursively enumerable *iff* it is generated by an EIL-system.

There are many ways to relate formal languages and machines. In a broad sense, the former can be understood as variations of a grammar +

alphabet system, while the latter have the Turing machine as a paradigm. A word is accepted by a Turing machine *iff* when fed as an input the machine halts. The following theorem relates machines and languages: A language is recursively enumerable *iff* it is accepted by some Turing machine *iff* it is generated by some (minimal) set of productions. In other words, if a language is generated by a grammar, there exists a Turing machine that will accept it, which can be proved using Church's thesis. Conversely, the inputs accepted by a Turing machine form a language with a certain grammar. This is not straightforward, but the corresponding grammar will consist mainly of the inverted productions of the machine, with an alphabet extended to include the internal states. It is possible to think of different automata when transitions other than the canonical Turing-like are allowed. Nevertheless, the transitions can always be described as a combination of the basic set, and this is all that is needed to achieve universal computation.

1.5 Computational time complexity

The Halting Problem demonstrates that there are unsolvable problems, which in the Turing model is equivalent to an automaton running for ever. Is there anything between the class of solvable and the class of unsolvable problems? When the computational resources are limited, one wants to know how much effort is needed to solve problems. One of these resources is time: does the existence of an algorithm guarantee that the time required to solve a problem

will be "reasonable"? In trying to answer this questions computer scientists developed the concept of *tractability*. A problem is tractable if the time needed to solve it is a polynomial function of the size of the problem. For example, an algorithm to sort an arbitrary set of numbers can be made to run in $O(n \log n)$, where n is the cardinality of the set; therefore, this algorithm is tractable because the time is bounded by $O(n^2)$. A problem is *intractable* if no algorithm running in polynomial time is known to solve it, i.e. any algorithm will run in super-polynomial time in the size of the problem. A famous intractable problem is the Traveling Salesman problem: a salesman wishes to make a tour, visiting a given number of cities exactly once and finishing at the city he starts from. There is a cost c_{ij} in traveling from city i to city j, so that every instance of the tour will have a total cost computed as the sum of c_{ij} for all cities in the tour, so that the problem is to find the tour with minimal cost. No algorithm is known to solve the Traveling Salesman in polynomial time. One can understand intuitively the reason: starting at the initial city, the space of possible solutions forms a tree growing factorially ($n!$, faster than exponentially) with each city, and which has to be explored in order to find the optimal solution.

These ideas are formalized in the theory of NP-completeness. The class P is defined as containing all problems for which a polynomial time algorithm exists, i.e. there exists a TM that halts for any input. The class NP is defined as the class of problems which are *verifiable* in polynomial time:

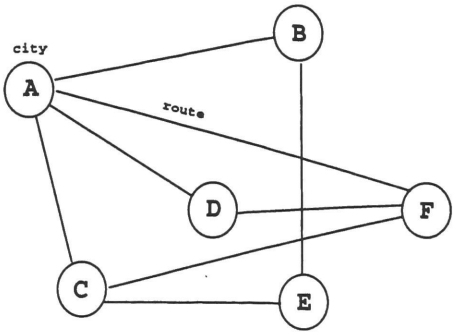

Figure 1.2: The Traveling Salesman problem

given a guessed solution, it takes polynomial time to verify whether it is a solution or not. It is easier to take the Hamiltonian Path problem as an example: in this case, there is not cost function between cities, they can be connected or not. The question to answer is whether there is at least one path that touches all the cities once and only once, starting at the first one and ending at the last one. One can imagine then, that a *non-deterministic* Turing Machine ($NDTM$, in which the transitions between states is defined by a probability distribution) generates a path at random in the first step of the computation, and then in the second step a *deterministic* one (DTM) verifies whether the path is a solution. Both steps are clearly polynomial in time, and therefore the Traveling Salesman problem belongs to the class NP.

The outstanding question as to whether $P = NP$ is important because if $P \neq NP$, then the problems in P are tractable, whereas those in $NP - P$ would be intractable. A basic idea in NP-completeness theory is that of *poly-*

nomial transformation. The solution to a problem can always be reduced to the question of whether the corresponding TM stops, which in its turn can be cast as a decision problem: if the TM stops, the answer is *yes*, and *not* otherwise. Two problems can be transformed polynomially if: (a) there exists a function F that transforms one instance of a problem into an instance of the other one, such that an answer 'yes' to the original problem implies an answer 'yes' to the problem transformed by F and vice versa; and (b) F can be computed in polynomial time. When two problems satisfy conditions (a) and (b) they are defined as *equivalent.* We can now define NP-complete problems: a problem is NP-complete (NPC) if it belongs to NP, and if it is equivalent to any other problem in NP. The class NPC might then be considered the most difficult problems in NP; moreover, the definition implies that if an algorithm is ever found to solve an NPC problem in polynomial time, *any* other problem in NP would become tractable. Several problems have been shown to be NPC, including the Traveling Salesman, the SAT problem (find an assignment of logical variables such that a given logical statement is true), as well as many problems in optimization theory, physics, neural networks, among others.

1.6 Algorithmic complexity

Time complexity is not the only theory of computational complexity. There is a growing interest in algorithmic complexity, based on the minimal descrip-

tion of an algorithm as a measure of its complexity [Li 93]. This measure of complexity is related to the concept of information content of an individual object. The origin of Shannon's information theory is the extension of thermodynamic entropy, which applies to states in an statistical ensemble, to ensembles of messages. Then information measures the average degree of "surprise" in the reception of a message, and is therefore a statistical quantity determined by the properties of the ensemble. But it is natural to define the information content or complexity of, say, a text or a picture, as the smallest description from which they can be reliably reconstructed. The problem that arises immediately is that the algorithmic complexity will depend on the *coding* strategy, that is to say, it is possible that the relative complexity of paintings by Picasso and by Dali changes with the description strategy.

Surprisingly, it was shown independently by Solomonoff and Kolmogoroff [Solomonoff 64, Kolmogorov 65] that the concept of algorithmic complexity can be formalized to avoid this problem. Some technical problems in the original formulation (for instance, the joint complexity was not strictly subadditive in the elements of the pair) were solved by the introduction of *prefix complexity* [Levin 74, Chaitin 75], resulting in the *Invariance Theorem for (prefix) complexity*. In order to understand the concept we need some definitions.

First, it is possible to map any enumeration of partial recursive functions $\phi_1, \phi_2...$ (as discussed in Section 1.3) into *partial recursive prefix functions*

ϕ_1', ϕ_2'..., so that the inputs for which the corresponding Turing Machines T_1', T_2'... halt are *prefix-free*: no input can be a proper prefix of any other one, and the inputs are binary strings. In other words, for any x, y in the set of inputs, x cannot occur as the initial substring of y. This is related to prefix codes in coding theory; these codes are uniquely decodable because the prefix-free condition avoids multiple interpretations of a given coding string (for further reading, see [Li 93]). The Turing Machines that correspond to this enumeration are called *self-delimiting*, because the input has to carry within itself information about when to halt - there are no special halting symbols in the tape, made up only of 0's and 1's. In the remainder of this section we will always refer to self-delimiting TM's.

Second, the conditional algorithmic complexity or algorithmic entropy of a binary sequence x given y respect to the Turing Machine $\mathring{\mathcal{A}}$ is defined as the size of the minimal program which outputs sequence x when y is fed into $\mathcal{A}$:

$$K_{\mathcal{A}}(x|y) = min\{|p| : \mathcal{A}(p, y) = x\}$$

where $|p| = \log_2(length(p))$ is size of the binary string that represents p in bits.

We can now formulate the Invariance Theorem: *There exists an Universal Turing Machine $\mathcal{U}$ such that for any other TM $\mathcal{U}'$*

$$K_{\mathcal{U}}(x|y) \leq K_{\mathcal{U}'}(x|y) + c_{\mathcal{U}'}$$

for all binary strings x,y.

That is to say, the conditional complexity is defined up to a constant that depends on the machine but not on the program itself, $|K_{\mathcal{U}}(x|y) - K_{\mathcal{U}'}(x|y)| \leq c_{\mathcal{U}\mathcal{U}'}$. Finally, one can define the "absolute" complexity as $K(x) = K(x|\lambda)$, by simply including the input y as part of the program p. The equal sign holds, of course, within an additive constant as prescribed by the invariance theorem. In the remaining of the section we will omit this constant from all equivalence and inequality relations.

Unfortunately, K turns out not to be a partial recursive function. Intuitively, the minimal description of an object, say a real number, is a numerable set. Then, one can apply the usual diagonalization argument to show that there are descriptions not in the initial enumeration. Nevertheless, under certain conditions its value can be bound or statistically estimated. For instance, the distribution of the number of strings of a fixed length as a function of the algorithmic complexity is given by $\mathcal{N}(K) \sim 2^K$, which implies that most strings have a complexity $\mathcal{K}(x) \sim |x|$. In particular for a typical natural number we have $\mathcal{K}(n) \sim \log_2(n)$. Moreover, K has several properties similar to Shannon's entropy, which form the basis of what is known as algorithmic information theory. One can define the *joint* complexity of a pair (x, y) as the minimal program which outputs x followed by y; it satisfies symmetry $K(x, y) = K(y, x)$, subadditivity $K(x, y) \leq K(x) + K(y)$, and the relationship between conditional and joint complexity $K(x, y) = K(x, y) - K(y)$.

Among the applications of algorithmic information, it has been shown

that the average algorithmic entropy of a member of a thermodynamic ensemble (i.e., the average complexity in describing, say, the positions and velocities of the molecules in a gas) is almost identical to its thermodynamic (or Boltzmann-Gibbs-Shannon) entropy [Bennett 82], opening the possibility of a thermodynamics algorithmically based (for a discussion, see [Li 93, Zurek 88]. It is evident that the possibilities of applying a consistent definition of descriptional complexity to biology are enormous, and still to be explored.

1.7 Formal and natural languages

As mentioned in the previous sections, an important factor in the development of the modern theory of formal language is the work of N. Chomsky on formalization of natural languages. In the Chomskian approach, the central problem of the theory of language is to explain how people can speak and understand new sentences, stressing the creative or generative aspect of language. He proposed that the study of language should distinguish between "deep structures" and "surface structures". Surface structures are the actual written or spoken utterances, whereas deep structures exist in the mind. These deep structures have strong regularities, shared amongst all human languages, and are the consequence of the capacity for digital computation of the brain. The study of the deep structures should reveal then the existence of an Universal Grammar, characterized by its formal properties, and

itself a product of certain unique areas of the brain dedicated to language processing. The Universal Grammar would explain then why languages from historically disconnected cultures can have common deep structures.

To understand the generative power of formal languages, Chomsky showed that they form a hierarchical structure, i.e. the complexity of sentences in a language is determined by the grammar that generates it. By the same token, there exists a hierarchy of machines, whose computational power determines the class of languages they can accept:

(i) A grammar is *regular* if every production is of the form $A \to bC$ or $A \to b$, where A and B are non-terminal and b is terminal. Regular languages can be accepted by *finite-state automata*, machines that have only a finite number of internal states and transitions between them, as opposed to the potentially infinite resources that Turing machines have through tape access.

(ii) A grammar is *context-free* if every production is of the form $A \to w$, where w is any word containing terminal and non-terminal symbols. In other words, the left side of the production consists of a single non-terminal symbol. Context-free languages are accepted by push-down automata, also known as stacks.

(iii) A grammar is *context-sensitive* if the restriction for context-free grammars on a single nonterminal symbol in the left side of the production is removed. The production are of the form $yAz \to ywz$, where y, w, z are words, and $w \neq \lambda$. Context-sensitive languages are accepted by *lin-*

ear bounded automata, which are Turing machines with the restriction that the read/write head cannot leave the tape space holding the original input, within which the head has to operate.

(iv) A grammar is *phrase structured* if it has no restrictions. Accordingly, they are accepted by non-restricted Turing machines.

If we call $\mathcal{L}_g$ the set of words or languages generated by a grammar g, the Chomsky hierarchy states that $\mathcal{L}_r \subseteq \mathcal{L}_{cf} \subseteq \mathcal{L}_{cs} \subseteq \mathcal{L}_{ps}$, where r stands for regular, *cf* context-free, *cs* context-sensitive, and *ps* phrase structured.

The immediate question is: where is natural language located in this hierarchy? In [Chomsky 59] it is shown that natural languages cannot be described as regular languages, and therefore they must be generated by a machine more powerful than a finite automaton. Much work has been done to show whether natural language is context-free, i.e., whether natural languages can be generated by push-down automata or more powerful machines, but this is still a matter of debate. Moreover, the initial hopes that the generative approach sparked, like the possibility of an universal automatic translator, have not been delivered, and seem rather distant. Nevertheless, the importance of this approach resides in pushing the wedding between natural and formal languages, and in probing how far this relationship can be held.

1.8 Formal languages and biology

The theory of computing machines and formal languages, and in particular Turing's work, seem geared toward a description of mechanical and electronic computing devices. Is it possible to apply the same concepts to biological systems? Can we learn anything from a computational view of living systems? There are several answers to these questions, depending on the specific biological context.

First, since Von Neumann's pioneering work on large-scale computing devices and formal descriptions of living systems [von Neumann 56, 66], there has been a two-fold interest in molecular computation: to reach the extreme limits of miniaturization and to understand the complex behavior of life in terms of a symbolic theory of the manipulation of information. Moreover, the discovery of the genetic code prompted many speculations on the formal nature of the genetic system [Leff 94]. In the particular case of DNA, it is evident that any system able to: (a) encode strings of any desired size, and to (b) operate on those strings in such a way that some combination of the minimal set of rules for grammars is fulfilled, will be capable of computing. The combination of DNA and DNA enzymes certainly satisfies these two basic requirements, and indeed a number of formalisms have been proposed to establish the computational power of DNA based computers [Paun 98], although actual implementations have been scarce so far (some of these issues are discussed in the Addenda).

Second, it has been proposed that the existence of logical gates implemented by enzymatic reactions can achieve universal computation, i.e., that it is possible to implement an universal Turing Machine where the tape and internal states are represented by the concentrations and activation and/or allosteric state of enzymes in a network [Hjelmfelt 92, Magnasco 97]. Enzymatic cascades in signal transduction, bacterial chemotaxis, the genetic program that leads to the consolidation of long-term potentiation, the development of an organism in general, can be thought of as instances of computer programs, and therefore should be studied with the same theoretical tools, in terms of computability, tractability, etc. [McAdams 95]

Third, the work of McCulloch and Pitts on neural networks showed that it is possible to achieve universal computation in a network of model neurons, based only on the thresholding nature of neural processing to implement logical functions, and on delayed connections to implement a dynamic memory (analogous to 'flip-flops' in electronics) [McCulloch 43]. In the case of neural networks, the connection with computer theory is more evident and necessary, from the simple decision making networks in the escape system of insects, to the modeling of semantic and syntactic networks in human speech.

In all three scenarios, the application of computational concepts has led to insight into the workings of biological complex systems. However, what will interest us is the interplay between the functional requirements of achieving robust and reliable computation, and the physical constraints imposed by

the capabilities of the actual computational device, be it oligonucleotides, enzymes or neurons, and a noisy environment, be it endoplasmic space in cells, excitable membranes in neurons, or neural networks in the brain. In pursuing this idea, we will encounter a number of questions: *how does noise affect biological computation*, or in other words, *how reliable can it be? can we program a biological computer? what are the computational strategies compatible with the resources available to biology?* We will try to answer these questions in the next chapters, as well as the new ones collected along the way.

Part II

Of messages and noise in the language of neurons

Chapter 2

A Role for Noise in the Generation of Action Potentials

Even the vagaries of chance have their place in Nature's scheme; that is, in the intricate tapestry of the ordinances of Providence... Moreover, what keeps the whole world in being is Change: not merely change of the basic elements, but also change of the larger formations they compose.

Marcus Aurelius

2.1 Introduction

The intra- and extra-cellular transmission of information must be robust and reliable processes. One of the solutions evolution has found for the extra-cellular transmission of information, both robust and extremely fast, is the generation of action potentials in the nervous system. As all cellular processes, the production of action potentials takes places in the presence of a thermal noise. In communication theory, noise is always the cause of a degradation of the transmitted messages, and the main goal in the design

of communication systems is to circumvent this problem. How is noise dealt with in biological systems? In this and the next chapter we will investigate the role of noise in the propagation of information by the nervous system, to show that more than a passive nuisance, it is inextricably linked to the transmitted information.

The basic mechanism for the production and propagation of action potentials is the following: at resting potential ($\sim$ -70 mV), the net ionic current between the positively charged outside and the negatively charged inside of the cell is zero. As the cell depolarizes (when synaptic input drives the potential to more positive values), the voltage-gated Na^+ channels open, causing an inward current of positive sodium ions. This Na^+ current causes more depolarization and therefore potentiates itself. At the same time, the depolarization spreads to the neighboring membrane, where the initial process is repeated and so the action potential propagates. After reaching $\sim$ 50 mV, the depolarization inactivates the Na^+ channels, and with some delay respect to the initial activation of the Na^+ channels, it activates the voltage-gated K^+. Potassium ions are in excess inside the cell, so that this opening causes an outward positive current; together with the inactivation of Na^+ channels this drives the membrane to its resting potential.

This description, first put forth by Hodgkin and Huxley in their seminal work on the giant axon of the squid [Hodgkin 52], implies that two basic requirements are needed for the regeneration of action potentials: the rec-

tification (ions can only flow in one direction) and the negative incremental resistance (during deactivation, the ionic flow decreases while the potential increases) for the ionic flow due to potential differences across the cell membrane. We will show that these phenomena occur naturally in Brownian transport as a consequence of a mechanism we term *entropic ratcheting*, by which flow obstruction is caused by entropic rather than energetic barriers. We will prove that negative incremental resistance and rectification cannot occur on 1D spaces like the circle or the line. We will construct an explicit two-dimensional model, and its collapse onto a branched 1D backbone, showing negative resistance and rectification as a consequence of a geometric symmetry breaking. Finally, we will discuss analogies and relevance to biological ion channels, in particular for channel inactivation and blocking [Cecchi 96].

2.2 Negative Incremental Resistance and Rectification

If a current arises as a result of some force, then it will flow "downhill", in the direction in which it dissipates energy into heat; thus, the resistance is always positive. The diminishing of a current as the driving force becomes stronger is called negative incremental resistance, or just plain "negative resistance" (NR); devices that display NR exist and have important technological applications; these devices are, usually, also rectifiers.

If energy is provided through suitable bias voltages, so as to displace the steady state of operation to the NR region, then the negative slope of the current can give rise to interesting instabilities like relaxation oscillations [Minorsky 62, Guckenheimer 83]. A typical such device from electronics is the tunnel diode. Much more importantly, the very existence of our nervous systems depends upon the ability to generate action potentials to transmit nerve impulses along axons; this phenomenon requires at least one NR rectifying device, which is known to be the sodium channel [Kornacke 69, Hodgkin 52, Hille 75a, Dubois 83, Alberts 89].

To understand how these phenomena can arise in the context of a cell, we will study a Brownian particle in a periodic potential, embedded in an equilibrium, constant temperature bath, and subject to a single force trying to advect it. This type of process is described by a Langevin equation of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \xi(t)$$

where $\mathbf{x}$ belongs to some Euclidean space, ξ is Gaussian white noise satisfying $\langle \xi_i(t)\xi_j(s)\rangle = 2kT\delta_{ij}\delta(t-s)$, and $\mathbf{f}$ is of the form $-\nabla V(\mathbf{x}) + \mathbf{F}$ with $\mathbf{F}$ constant, i.e., $\mathbf{f}$ is a vector field independent of time. We will assume V to be periodic along the direction of $\mathbf{F}$.

Why would it be interesting to search for NR in transport processes of this form? First of all, these processes have very wide applications [Risken 89]. Second, transport in symmetry broken potentials ("ratchets") [Risken 89,

Adjari 93, Magnasco 93, Magnasco 94, Millonas 94, Astumian 94, Doering 94] has been shown to be analogous to conduction through electronic diodes, capable of rectification; however, the mechanism through which a *tunnel* diode provides NR is intrinsically quantum mechanical; hence, it would be interesting to provide classical analog. Finally, ratchet potentials in one dimension can be shown not to have NR, as we will now do.

In one dimension, an equation of the form

$$\dot{x} = f(x) + \xi(t), \qquad \langle \xi(t)\xi(s) \rangle = 2kT\delta(t-s)$$

has an associated stationary Fokker-Planck equation

$$\partial_t P(x,t) + \partial_x J(x,t) = 0$$

$$J(x,t) = fP(x,t) - kT\partial_x P(x,t)$$

whose steady state, for periodic $f(x)$, can always be solved in quadratures. If f has a zero spatial average, then it is the gradient of a periodic potential, and detailed balance and Boltzmann weights hold. If f does not have a zero average, then it can be written as $f = -\partial_x V + F$ with V periodic, and the stationary state can be solved in double quadratures [Magnasco 94]. The Fokker-Planck probability current J as a function of F is given by:

$$J(F) = \frac{kT(e_{kT}^{2\pi F} - 1)}{Q(F)} \tag{2.1}$$

$$Q(F) = \oint\oint e_{kT}^{V(x')-V(x)+F(x-x')+2\pi F\Theta(x'-x)} dx'dx \tag{2.2}$$

where e^x_{kT} means $\exp(x/kT)$. Q is positive; moreover, $\partial_F Q$ is also positive, and satisfies $kT\partial_F Q < 2\pi Q$. This inequality implies $dJ/dF > 0$ and hence Brownian transport in one-dimensional periodic potentials (under a steady force and in white noise) cannot show NR. (We will show later that this only holds if the underlying space is topologically trivial).

2.3 An Example in Two Dimensions

Now we will construct an explicit example in two dimensions. First we illustrate the notion of entropic barriers, through a very simple example. Let us consider the following potential

$$V(x, y) = x^2 \left(\cos(y) + 1.1\right)$$

We see that the potential is identically zero on the y axis, and bigger than zero everywhere else. Hence there are no true energy barriers impeding motion of a Brownian particle along the periodic direction y. However, the shape of the potential around the y axis is also important at nonzero temperature. For any given y, a slice of the potential along that value of y is a parabola; however, as motion progresses along y this parabola opens and closes periodically. In the absence of an external force, a Brownian particle will spend more time around $y = \pi$ than on the bottleneck $y = 0$, every now and then jumping one period up or down in y, as if there actually were an energy barrier; these are called *entropic barriers* because, unlike a true activation energy, the time scales they induce do not follow Arrhenius-Kramers laws. In other words, as

we will see, the violation of these laws is responsible for the non-monotonic behavior of the current as a function of the applied force F.

We will now break the parity symmetry along the y axis; if we loosely call any parity broken potential a "ratchet", we can say the following potential is an *entropic ratchet:*

$$V(x,y) = x^2 \left(\cos(y + \ln\cosh x) + 1.1\right)/2 \tag{2.3}$$

where $\ln\cosh x$ is just an easy way to make a function that is both even and linear in x for large x. The equipotentials are now symmetry broken and look like a herringbone pattern, see Fig. 2.1. Our dynamics will be given by

$$\dot{\mathbf{x}} = -\nabla V + F\hat{\mathbf{y}} + \xi(t) \tag{2.4}$$

where $\hat{\mathbf{y}}$ is the unit vector along y, and the noise correlators are as before.

So, if we apply a force F along y, if the force is positive the particle will move forward without problem. However, if the force is negative, the particle will move backwards, but every now and then it will get into a spine of the herringbone. It will progress upon the spine for a while, deeper the stronger the force, and then will have to go *against* the force in order to climb back up and get again on the backbone. Thus it will be locked for a while, because the energy required to climb back up is an honest activation energy, and the time required to do so obeys an Arrhenius law. But the probability that a particle on the backbone will get into the spine does not depend exponentially on F, and so the particle spends a larger proportion of time blocked away in

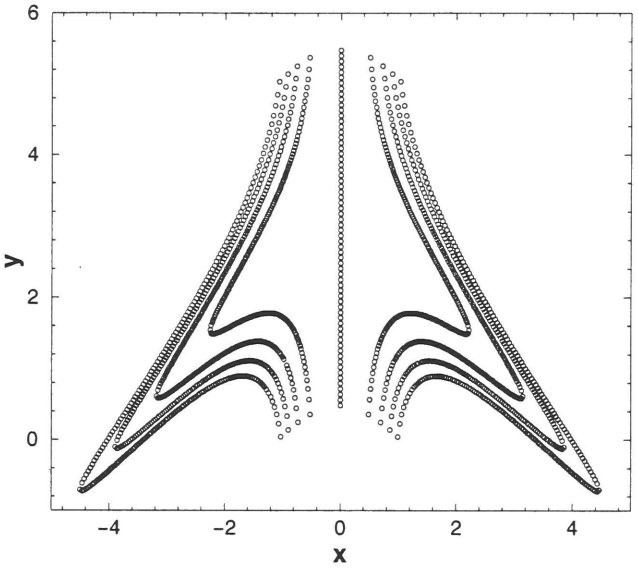

Figure 2.1: The equipotentials of the model.

the spines as F becomes more negative. The overall time scale to get into and out the spine is not Arrhenius-like, yet near the bottleneck the potential energy of points in the spine and in the backbone is the same. If it was, the time the particle spends stuck in the spine would cancel out with the time it spends in the backbone, and the current would be a monotonic function of F.

We have evolved numerically Equation 2.4 to obtain the average speed of the particle as a function of F and kT. The mean speed equals the Fokker-Planck current times the period (2π) of the potential. The result of our simulations is shown in Figure 2.2.

In performing a numerical integration of Equation 2.4, one encounters several numerical problems. The simplest method for numerical integration is the Euler method,

$$x(t + \Delta t) = x(t) + f(x)\Delta t + \sqrt{2kT\Delta t}\eta$$

where the η are random gaussian numbers with unit variance. While the Euler method is first order in Δt for an ordinary differential equation, it is only one-half-order for a Langevin equation, a property characteristic of diffusion. This implies that extremely small time steps have to be used for accurate integration. Several methods have been proposed for increasing the order of the integration, including stochastic versions of the Runge-Kutta algorithm [Honeycutt 92]. However, the spines in our potential are a *stiff* problem, being long and skinny, which would limit the applicability of a

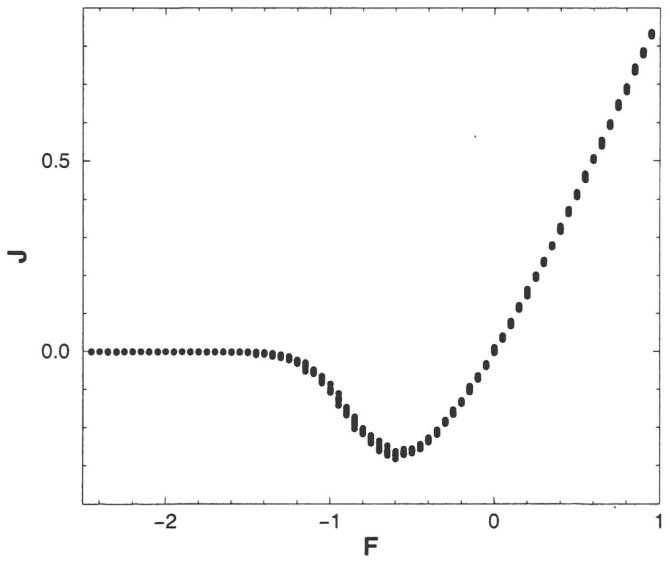

Figure 2.2: $J(F)$ at $kT = 1$ as computed from Langevin simulations. Each run lasted for 10^5 units of time; ten runs were done for each value of F.

Runge-Kutta scheme even in the absence of noise. The reason is that the time scale has to be small in comparison with the relaxation time on the fastest direction, in this case the skinny direction normal to the axis of the spine. But then the interesting time evolution is that *along* the axis of the spines, which becomes painfully slow.

We have devised a method to cope with this problem. The fundamental problem in developing a Runge-Kutta scheme is that somehow one is assuming analyticity both of the vector field and of the solution; the last one is just not there. But one can safely assume analyticity of the vector field alone. If we expand the latter to first order,

$$\mathbf{f}(\mathbf{x} + \mathbf{\Delta x}) \approx \mathbf{f}(\mathbf{x}) + \mathbf{\Delta x} \cdot (\nabla \mathbf{f})(\mathbf{x})$$

We use the fact that $\mathbf{f}$ is curl-free (and thus $\nabla \mathbf{f}$ is a symmetric tensor) to diagonalize it and rotate to its eigenbasis. In this base, the problem locally becomes a cross-product of independent Ornstein-Uhlenbeck processes. Thus, the question is: if we know we currently are at position x at time t, what is the probability that we will be at position x' at time t'? The answer to this question is the Fokker-Planck propagator (Green function), which is known analytically for the Ornstein-Uhlenbeck process [Risken 89]. It is a Gaussian, centered at the position x' that would be the solution for the deterministic case at time t', and with a width which has been changed because of the compression/expansion due to $f' \equiv df/dx$. Thus it is a trivial matter to generate a new value x' with the correct probability distribution for an

Ornstein-Uhlenbeck process; our numerical method then reads

$$x' = (x - x_c)e^{-f'\Delta t} + x_c + \left((kT/f')(1 - e^{-2f'\Delta t})\right)^{1/2} \eta$$

$$x_c = x - f/f', \qquad \Delta t = t' - t$$

for each eigendirection of the Hessian. The new coordinates are then rotated back to the original frame. The advantages of this method over a stochastic Runge-Kutta procedure are two. First, it solves *exactly* the Ornstein-Uhlenbeck process, by construction, and hence any linear problem. Second, it can handle stiff problems more easily; it will not lose accuracy if the time step is larger than the relaxation time scale of the fast direction, because it will not *overshoot* and generate dynamical instabilities; one can concentrate on the more interesting slow time scale. The disadvantages are also two: analytic knowledge of the Hessian is required, and a matrix eigensystem calculation has to be performed, together with two coordinate frame transformations. Thus our method rapidly loses ground to the stochastic Runge-Kutta method for high dimensionalities, unless the problem is quite stiff or the Hessian is sparse. For our stiff problem in two dimensions, this method is extremely well adapted; Fig. 2.2 was generated in a few hours of CPU time of a workstation and shows almost no trace of the residual noise typical of direct Langevin simulations.

2.4 Collapsing the Dynamics

Back to our problem, we would like to get some more explicit understanding of the $J(F)$ curve we have just obtained numerically. Let us consider a one-dimensional cartoon of the system: we can represent the backbone as a circle, with a single spine of length L attached to it, at an angle θ, as in Fig. 2.3a. There is no potential, just the steady force. At the vertex, we have to impose that the probability be continuous: the limit of $P(x)$ as we approach the vertex from all *three* sides must be the same, P_{vert}. The current should be conserved at the vertex, so that probability does not accumulate. Then, the probability density on the circle is constant, and so is the current on the backbone: $J = FP_{vert}$. The current on the spine should vanish, and hence the probability along the spine $P(x) \approx \exp(-Fx/\tan(\theta))$. Therefore, the final current is

$$\frac{F}{J(F)} = 2\pi + \frac{kT}{F\cos\theta}(1 - e^{-FL\cos\theta}_{kT}) \tag{2.5}$$

which has the right qualitative form (Fig 2.3b), except for a single detail. For F large and positive both cartoon and full case converge algebraically to unity. But for F large and negative, $J \to 0$ exponentially, while in the full case it seems to die faster. This behavior can be understood if we recall that the spines are truly different: first, there is a potential along them; second, they are not of length L, but rather arbitrarily deep. The center of the spines lies approximately at $y = \pi - \ln\cosh x$, and the potential there is

$V = 0.05 * x^2$. The current in this case equals

$$\frac{F}{J(F)} = 2\pi + 2\int_0^\infty \sqrt{1 + \tanh^2(x)}e^{-0.05x^2 + F\ln\cosh(x)}dx \qquad (2.6)$$

Thus, for a large and negative force, the particle entering the spine will find a stable fixed point at $x \approx 10F$. Thus the effective length of the spine *increases* as F becomes more negative; the activation barrier is the product of this effective length and F, and hence *quadratic* in F rather than linear. So we have shown that NR is actually possible in one dimension, but only if the topology of the space is more complicated than just a circle.

2.5 Conclusions

We have presented an explicit example of a Brownian transport process showing NR. The NR region is generated through an essentially entropic process. There are no obstacles, no energy barriers to the motion of the particle along the y axis. There is just a finite probability of exploring space a bit out of the center, and of ending up blocked inside a spine. Since we are in a thermal, rather than quantum, situation, our system has much more similarity to biological ion channels than to electronics. There are two well known instances in channels where NR is observed. The first one is in channels that can have an *inactive* state; this is the case of the Na channel, but some K channels also have this property. This is extremely important biologically, because the NR resulting from inactivation is essential to the regeneration of action potentials. The second one is the blocking of channels

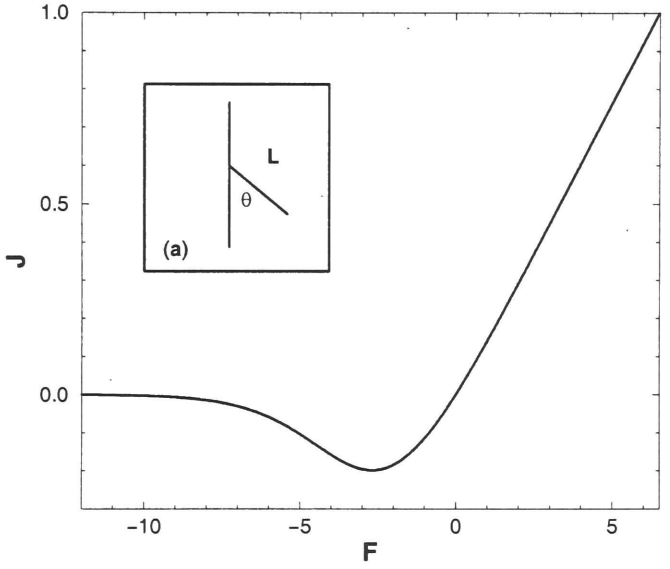

Figure 2.3: (a) The phase space for the cartoon. (b) $J(F)$ as given by Eq. 4

through large ions; this is important experimentally, because in order to assess properties of new channels, biophysicists will test for changes in behavior when the channel is "poisoned" with various compounds of known effect on known channels. The standard poison arsenal includes several large ions, that can get partially into the channel and block it; for example, tetraethylammonium is used for Na channels [Armstrong 65] and Mg [Hille 78], Cs [Hagiwaram 76] or polyamines [Lopatin 94] for K channels. It is worth noting that the I-V curves of such channels look extremely similar to that of our model, and, furthermore, several of the experimental measurements [Dubois 83, Armstrong 65, Hagiwaram 76] show the faster-than-exponential decay of our model, while standard theoretical models with *fixed* barriers [Hille 78, Hille 84] show exponential decays; this discrepancy can only be solved through models having barriers that depend on the field. This might mean that the large ion buries itself deeper and deeper into the crevice of the channel, getting more and more stuck and having to climb a larger distance against the potential to get out, just as for our spines. There are also similarities with the case of channels that display inactivation. These channels are hypothesized to have three states: closed, open, and inactive. The transition between open and inactive states has been modeled with a "ball and thread" mechanism [Alberts 89, Hille 84, Armstrong 77]: some mobile, charged, globular piece of the channel loosely attached through a long thread can get stuck in the mouth of the channel and block it; pretty much like a

large ion blocker. Our model does not include a closed state, since the central backbone is always "open"; thus we do not see the current going to zero as F becomes large and positive. The closed state can be bypassed by looking at the peak sodium current from a pulse rather than the stationary current [Hille 75b], and in this case the observed peak I-V curve is qualitatively similar to what we observe.

Chapter 3

Timing of action potentials: stammerers or euphonic couriers?

One must interrelate the languages of various peoples, and one should not make too many leaps from one nation to another remote one unless there is sound confirming evidence, especially evidence provided by intervening peoples. In general, one should put no trust in etymologies unless there is a great deal of concurrent evidence.

G.W. Leibniz, New Essays

3.1 The code of the single neuron

What is the language of a neuron? There have been many attempts to characterize neuronal codes in terms of temporal relations between spikes generated by the same or different neurons, in contrast with the more parsimonious hypothesis of rate coding. The evidence for the rate theory is well established in different brain areas, whereas the existence of a generalized temporal code has been more elusive. It is evident that the brain solves

many tasks with exquisite timing resolution, among others sound processing (think, for instance, of a piano concertist). Indeed, one of the most clear demonstrations of this has been found in the owl's sound localization system [Carr 88], where a network of delay lines implements coincidence detectors to compute inter-aural time differences. Beyond coincidence detection, all evidence for a complex temporal code, i.e. non-trivial structures in the spike train, appears to be indirect.

The temporal structure of neural responses has been related to the putative role of the synchronization of spike production and the oscillation of large neuronal ensembles. These features, which seem to be complementary, have been linked to a variety of behavioral and functional aspects of neural processing. The extent of the literature in this regard precludes any exhaustive systematization; nevertheless, we can mention some of the hypothesis for a functionality of synchronization and oscillation: *(i)* the solution to the binding problem in the visual pathway [Roelfsema 98]: cells coding for different features would be "bound" in an ensemble representing the whole object by engaging in a collective oscillatory state; *(ii)* the discrimination of stimuli in the olfactory pathway by synchronizing specific neuronal assemblies during the dynamic response to different odors [Laurent 96]; and *(iii)* encoding the onset of exploratory behavior across several brain structures [Nicolelis 95]. Audaciously enough, oscillations have also been hypothesized as the neural correlate of consciousness [Kahn 97]. Finally, regarding specific mechanisms,

it has been proposed that oscillations subserve an encoding strategy based on the relative phase of individual spikes respect to the collective oscillation, as a means to achieve more computational power in realistic continuous-time neural networks [Hopfield 95]. An alternative view postulates that oscillations are the consequence of a universal mechanism to build neural maps in a self-organizing way [Kohonen 96]. It requires a winner-take-all strategy for stimulus discrimination, implemented by lateral inhibitory connections; the existence of delays in these connections imply oscillations during the winner selection phase [1].

A different scenario in which the temporal structure of spike trains might play a role is the regulation of neuronal genomic expression. There is increasing evidence suggesting that plastic changes in neurons are triggered by specific patterns of electrophysiological activity, but the precise form this activity should take is not understood. In this regard, one can hypothesize that specific temporal structures could be needed to trigger enzymatic cascades leading to genomic regulation, or changes in the internal state of neurons other than electrophysiological, necessarily in a longer time scale. Evidence supporting this idea has been recently reported [Dolmetsch 97, De Koninck 97, Dolmetsch 98].

A criticism that the temporal coding theory has received is that neurons cannot rely on fine details of the structure of spike trains due to the

[1] Oscillations arise naturally and robustly in simulations of integrate-and-fire networks with delayed lateral inhibition (unpublished results)

noisy nature of spike generation [Barlow 72]. In this view, only averaging the number of spikes over a long time window, or eventually over a neuronal population, can be related to information transmission. On the other extreme, some neurobiologists suggest that part of what is usually discarded as "noise" in neural traces can be indeed signal, although not necessarily related to a simple input-output relationship with the stimulus, but rather the signature of an ongoing activity intrinsic to brain function [Arieli 96]. However, any temporal coding strategy (even within the last, more permissive, framework) implies a notion of *precision*, i.e., the relative (within the train) or absolute (with respect to a stimulus or a task) position of spikes must be tightly timed. In this regard, one of the issues that has been somehow overlooked in the debate is the interplay between the coding strategy and the particular stimulus or message being relayed by a neuron or neuronal ensemble. In functional terms, one would like to know whether the encoding strategy is dependent on the message, or whether eventually different messages can be relayed using alternate encoding modalities [Nicolelis 98].

In the next section we will show that the variability in the response (i.e., the precision of action potential timing) of a very general class of model neurons depends on the particulars of the input that drives them. We will also present experimental results supporting the predictions, which suggest that a simplification usually made in the application of information theory to neural processing is violated: noise is not always independent of the message

[Cecchi 99].

3.2 Noise in neurons is message-dependent

Brains represent signals by sequences of identical action potentials or spikes [Adrian 28], typically occurring in an irregular temporal pattern. Bryant and Segundo [Bryant 76], first noticed (in the mollusk Aplysia) that such neuronal variability (i.e., the precision of action potential timing) depends on the particulars of the input that drives them. This intriguing property has received renewed attention, notably among those searching for experimental evidence supporting different theories of neural coding [Mainen 95, de Ruyter 97]. Two representative reports sparking the recent revival of interest are those of Mainen and Sejnowski[Mainen 95] and de Ruyter van Steveninck et al.[de Ruyter 97]. In [Mainen 95], when the response of pyramidal neurons to injected current was studied, it was found, that the temporal pattern of firing is very irregular when the injected current is constant, but when the input current is "noisy", or has high frequency components, the resulting pattern is highly reliable. In the other report[de Ruyter 97], similar results were obtained from the H1 neuron in the visual system of the fly. In vivo recordings were made while stimulating visually with a pattern of moving bars. In the experiment, the pattern was moved horizontally, either at very small constant velocity (i.e., "quasi-static" condition for this motion-sensitive neuron) or with a computer-generated random velocity waveform.

It was found that the quasi-static condition produces irregular spike firing patterns, but more natural, time-dependent input signals yielded patterns of spikes much more reproducible, both in terms of timing and of counting precision. Thus, similar observations regarding timing precision of the spikes can be made in very different types of neurons under very different conditions.

3.3 Stochastic Integrate-and-Fire Model Neuron

It is straightforward to replicate the essence of the phenomena in question. For instance, one can use the simplest model of a spiking element, the leaky integrate-and-fire (I-F) model [Knight 72], which assumes the neuron as being a (leaky) capacitor driven by a current simulating the actual synaptic inputs. Membrane fluctuations are added, representing several internal sources of noise (cluttering of ion channels, synaptic noise, etc.). The system is described by the following Langevin [Risken 89] equation:

$$C\dot{V} = -gV + I(t) + \xi(t) \tag{3.1}$$

where C is the cell capacitance, V the membrane potential, gV the leakage term (g is a conductance), $I(t)$ is the input current, and $\xi(t)$ is Gaussian noise, $\langle \xi(t) \rangle = 0$ with autocorrelation $\langle \xi(t)\xi(t') \rangle = \sigma\delta(t - t')$ and such that when the potential reaches the threshold V_0 an action potential is generated and

the system returns to the equilibrium potential V_e, set arbitrarily to zero.

Fig. 3.1 shows the results of a numerical simulation of Eq. (3.1) in response to two different signals (in both cases there is noise). The left panels represent the (extreme) case of a time independent input, while the right panels are representative of more naturalistic inputs where there is high frequency components. For the former case one can observe a large scatter of the spike timings under successive trials. Note that this scatter becomes more pronounced as time progresses from the beginning of the stimulus. In contrast, when $I(t)$ contains high frequency components (right side panels of Fig. 3.1) spikes are clustered more tightly to the upward-going instances of the input and overall there is less trial-to-trial variability than in the time-independent case. The phenomenology described in Fig. 3.1 captures the main claim that renewed the interest in the subject; i.e., that apparently more "noisy" (in the misleading sense of "randomly fluctuating") inputs generate less variable responses.

We will offer a simple argument to explain the essence of the phenomenon of variability. Two important points are here clarified: (I) which is the most basic mechanism producing such variability seen in disparate types of neurons? and, (II) what (if any) are the implications for brain function? Regarding the first we will show that a simple geometrical argument can be used to explain the described phenomena for all generic neuron models. Concerning the second point, our results suggest that a fundamental notion

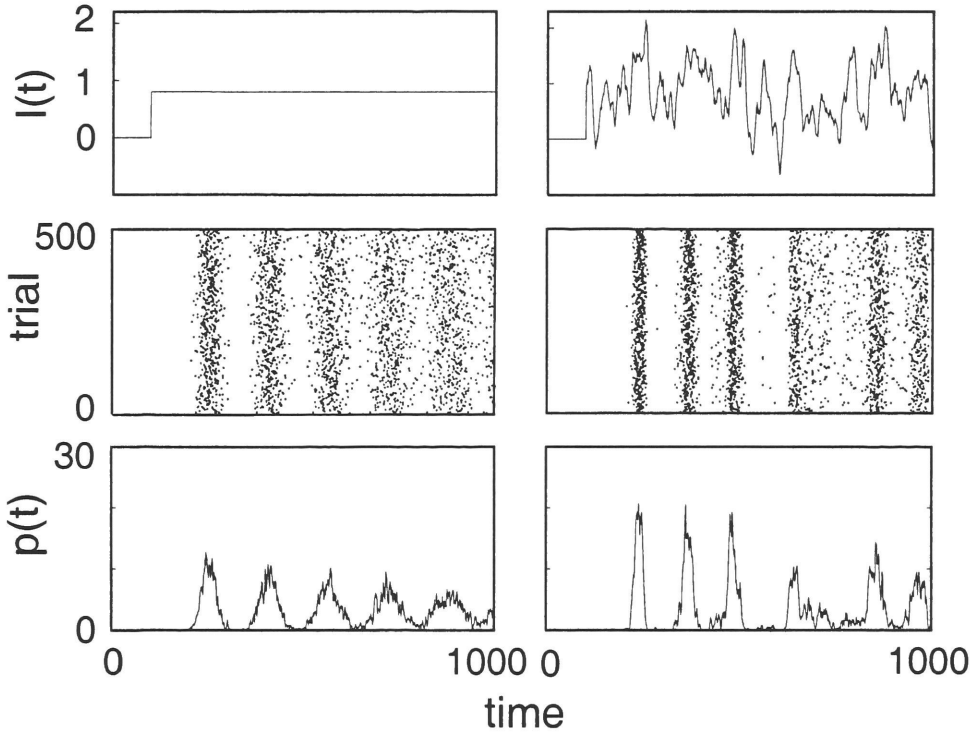

Figure 3.1: Rapidly fluctuating (panels on the right) inputs produce more reliable output spikes than a time independent stimuli (panels on the left). Panels in the top correspond to input stimulus - $I(t)$ in Eq. 1; middle ones to the 500 responses to the stimulus where the occurrence of each spike is denoted with a dot; and bottom panels depict the peri-stimulus time histogram describing the rate - $p(t)$ - at which spikes are generated in response to the stimulus $I(t)$.

of information theory (that noise and message are independent variables) will not be obeyed by neurons , unless there is a mechanism ensuring that most spikes are generated by relatively fast threshold-crossings.

In neurons, noise will affect in particular the membrane potential. In order to translate that into time noise, we have to scale the potential noise by something having units of *time/potential*, which is nothing but the inverse of the slope, or rising time of the potential. It is very clear then that an input signal having fast components will make the neuron membrane potential rise with a bigger slope (although this is not direct), and therefore, from a simple geometrical interpretation, we can see that in the presence of uniform noise the fluctuation of the spiking times should be smaller for fast inputs.

In a rough approximation, for an input with time constant τ, the deterministic evolution of the potential is $V \sim t^2/\tau$, which implies that near the threshold it scales as $\mathrm{d}V/\mathrm{dt} \sim 1/\sqrt{\tau}$, and so the timing imprecision goes like $\sqrt{\tau}$, in agreement with the results reported in [Mainen 95].

Thus, *the faster the voltage approaches the threshold from below, the more reliable the timing of the spike will be.* We proceed now to estimate this by using similar numerics than for Fig. 3.1. A rapidly fluctuating signal was constructed and the times at which each spike occurred were determined first for the deterministic condition (i.e., $\sigma = 0$), and subsequently for several hundred stochastic realizations (i.e., with identical $I(t)$ but different stochastic realization of the noise term in Eq. 1). At the same time the voltage deriva-

tive preceding each spike was computed over the last fifty time steps. The fluctuating character of the $I(t)$ used here (similar to that illustrated in the top right panel of Fig. 3.1) offers the opportunity to explore a wide range of $\mathrm{d}V/\mathrm{dt}$ crossings. To estimate the spike timing precision we define an index of "temporal jitter" as follow: it is the absolute time difference between the time at which each spike is generated in the noise-free simulation with respect to the time at which the same spike is generated in a stochastic realization. This quantity is averaged over all spikes and all realizations and the results are plotted in Fig. 3.2. It seems that the intuition behind our geometrical argument is essentially correct. The data points in Fig. 3.2 are scattered around the predicted (dotted lines) inverse relationship between the temporal jitter of the spike and the speed at which the voltage crossed the threshold. The trajectories plotted in both insets help to visualize the geometrical argument already discussed: inputs that raise more rapidly the membrane potential to threshold will produce less variable spike timing in presence of constant noise.

3.4 Signal-dependent noise is ubiquitous

The same basic phenomenon, of course, will affect any model of spike generation or excitable media, in which the noise in the voltage-like variable is translated into jitter of its threshold crossings. To provide for a specific example, we choose a widely used model of excitable media with continu-

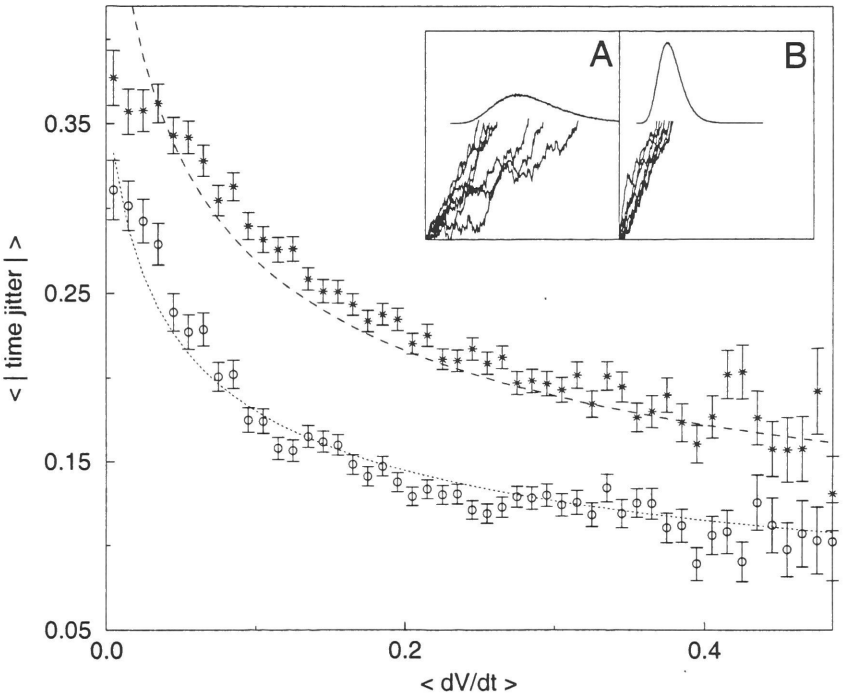

Figure 3.2: Computed mean spike temporal jitter (and SEM) as a function of the estimated membrane potential's dV/dt. Estimates are plotted for two noise variances: $\sigma = 0.01$ ($\circ$) and $\sigma = 0.02$ ($\star$). The expected inverse-law fit (with exponent ~ 0.35) is depicted with dotted lines. The examples in the insets illustrate two density distributions of spiking times resulting from relatively slow (inset A) or relatively fast (inset B) dV/dt threshold crossings. For each case, below each histogram, a few typical trajectories are also plotted showing the membrane potential preceding the threshold crossing.

ous dynamics, the Fitzhugh-Nagumo model (FHN) with additive stochastic forcing [Longtin 98], described by the following Langevin system:

$$\dot{V} = V - V^3/3 - W + I(t) + \xi(t) \tag{3.2}$$

$$\dot{W} = \phi(V + a - bW) \tag{3.3}$$

The variable V is the fast voltage-like variable, W is the *slow* recovery variable, and $\xi(t)$ is a zero-mean Gaussian white noise of intensity D. After a firing, the recovery produces an absolute refractory time T_R during which a second firing cannot occur, followed by a longer relative refractory time during which firing requires stronger perturbations.

It is biologically reasonable to assume that noise affects more the voltage-like variable. Thus, according to our argument, a high steady input will result in a faster approach to the threshold crossing and therefore reduce the probability of an untimely crossing. This is confirmed in Fig. 3.3, where the distribution of periods for two input amplitudes are plotted. Note the large change in variance (1.5 to 2.34, 56%) for comparably small change in the average period (36.4 to 39.1, 7%) in going from high to low input, and that the low input distribution has a supra-exponential tail. For comparison, the distribution of inter-spike intervals corresponding to the I-F model with constant input is shown. The distribution for high input tends to a Gaussian, while the low input has an exponential tail as expected for Poisson-like

statistics [2]. These results show the generality of the phenomenon.

Thus, in neural data, one should expect generically that the output noise depends on the input, i.e., there is a *message-dependent noise*. Experiments discussed next can gauge up the extent of its relevance for brain function. Recalling that an usual assumption made in the application of information theory to neurons is that the noise introduced by the communication channel is independent of the message being transmitted, we decided to explore to what extent this assumption is valid in real neural data. This is presented in Fig. 3.4, where different visual stimuli (i.e., the four messages) were presented to an anesthetized cat while electrophysiological recordings were made in the lateral geniculate nucleus (LGN), the second stage of processing in the visual pathway [3].

Raster plots in Panel A in Fig. 3.4 show the response to a moving bar with high contrast and high speed, and in Panel B the responses to a low contrast bar moving at low speed, in a window of 200 msec centered at the peak of

[2]n Fig. 3.1, Eq. 1 was numerically solved with parameters: $\sigma = 0.01$, $V_0 = 0.3$; steady input $I = 0.8$, fluctuating input was generated as an Ornstein-Uhlenbeck process with variance 0.01 and correlation time 0.1. (for easy of presentation $I(t)$ is plot as a ten points running average). In Fig. 3.3, Eq. 1 parameters are: noise variance 0.003, g=0.1, high input I=1.0, low I=0.1; the parameters for Eq. 2-3 are: noise variance 0.003, high input $I = 1$, low $I = 0.5$, $\phi = 0.08$, $a = 0.7$, and $b = 0.8$.

[3]Experimental methods as in [Martínez 98]. Briefly, cats (2.5-3 kg) were anesthetized with ketamine (10 mg/kg) and thiopental sodium (20 mg/kg; maintenance 2 mg/kg/hr, IV) and paralyzed with Norcuron (0.2 mg/kg/hr, IV). Temperature (37.5o-38o C), EKG, EEG, and expired CO2 was monitored throughout the experiment. Pupils were dilated with 1% atropine sulfate and the nictitacting membranes retracted with 10% phenylephrine. A multi-electrode matrix was used for the cortical recordings. Recorded signals were collected and analyzed by a computer running Datawave Systems software (Broomfield, CO).

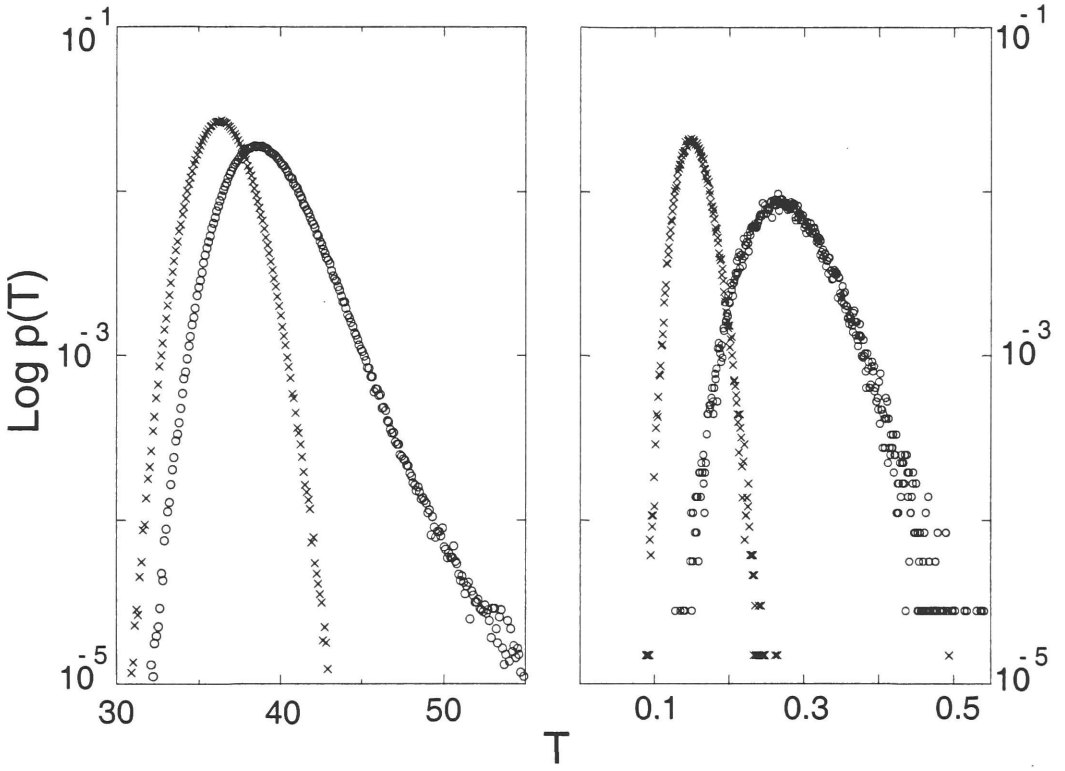

Figure 3.3: First passage time distributions for the FHN (left panel) and I-F (right panel) models, forced with constant input. In both panels, (×) corresponds to high input and (○) to low.

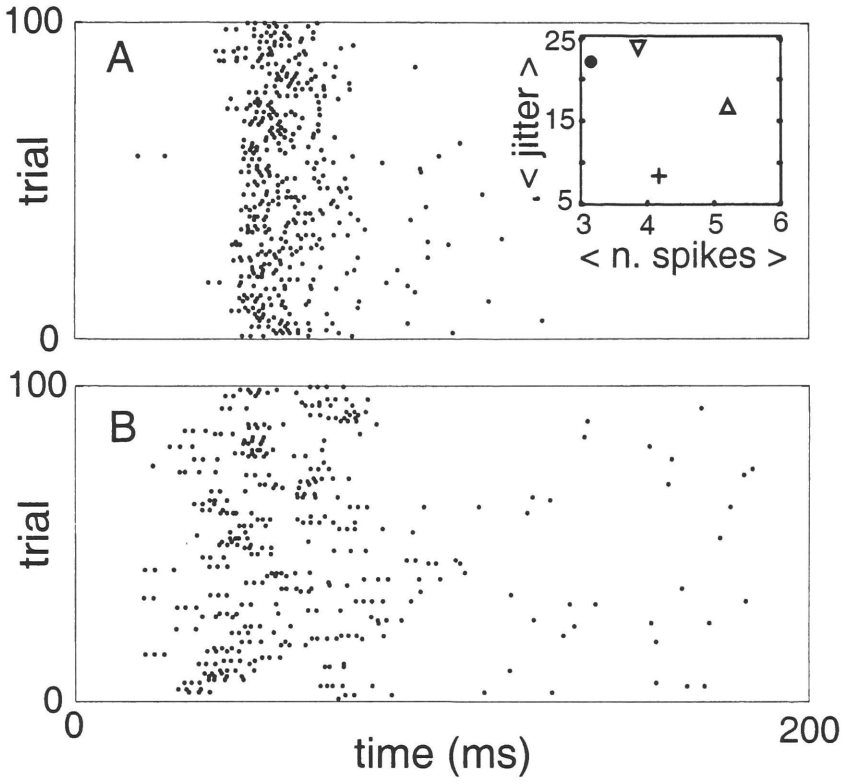

Figure 3.4: Experimental example showing that different "messages" elicit different noise (i.e., temporal jitter of spikes). Results gathered from recordings in cat's LGN where neuronal action potentials are recorded in response to the presentation of a moving bar with different contrast and speed. The inset in Panel A shows, for the conditions tested, the temporal jitter versus average number of spikes (collected during the window of 200 ms); ($\bullet$) corresponds to low contrast/high speed, ($+$) to high contrast/high speed, ($\triangledown$) to low contrast/low speed, and ($\triangle$) to high contrast/low speed. Notice the data from moving bars of different contrast originating very different jitter. The raw data for two of the experimental conditions are presented in the raster plots of Panels A and B. Responses obtained with a moving bar of low contrast moving at low speed are plotted in Panel B. The (less variable) responses obtained with a moving bar of high contrast moving at high speed are depicted in Panel A.

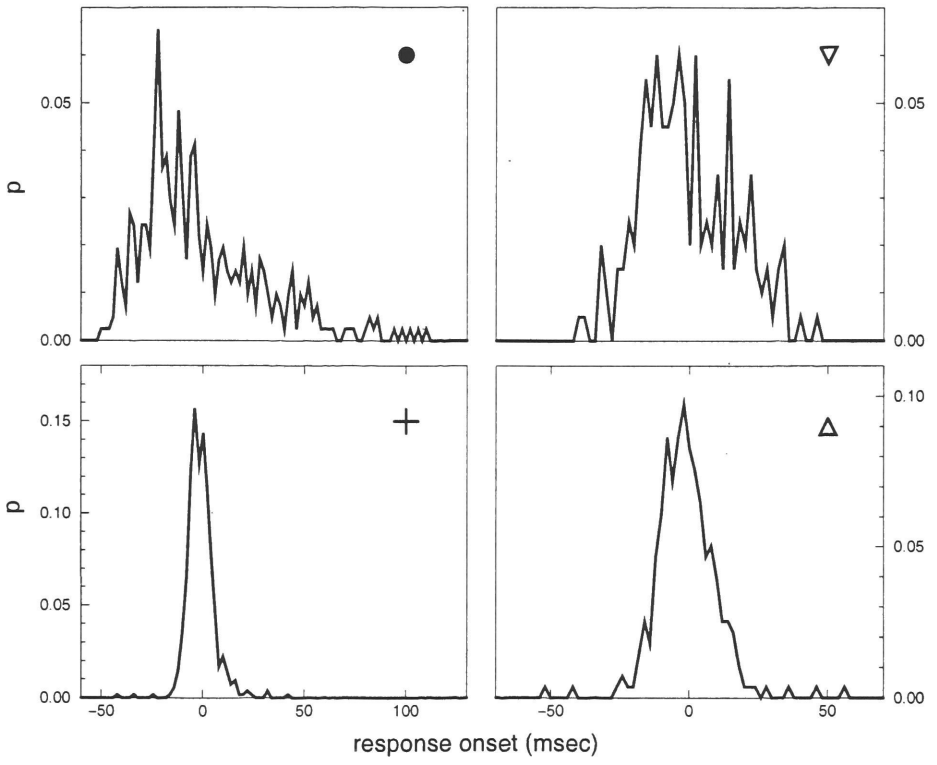

Figure 3.5: Probability distribution of response onset for the four stimulus conditions (1 msec bins). Symbols as in Fig. 3.4

the ON response. The responses in Panel A clearly display a higher temporal precision than those for the condition in Panel B. In the inset of Fig. 3.4 we quantify the noise for the four messages as the "jitter" of the response onset, defined as the average of the absolute value of the position of the first spike in each trial with respect to the mean. The probability distributions of the response onset are shown in Fig. 3.5. The messages include two conditions for the velocity and two for the contrast. For each velocity condition, the high contrast is more reliable than the low one: 8.4 msec vs. 22.1 msec, and 16.8 msec vs. 23.8 msec, according with the difference in number of spikes: 4.17 vs. 3.15, and 5.2 vs. 3.86. Note however that the high contrast/high speed and low contrast/low speed stimuli, although having a very similar spike response (4.17 vs. 3.86) have the biggest difference in jitter (8.4 msec vs. 23.8 msec). Alternatively, we can measure the entropy difference between the corresponding first spike probability distributions (1 msec bins), which varies from 2.0 bits for low contrast/high speed vs. high contrast/high speed, to 0.6 bits for high contrast/low speed vs. high contrast/high speed. Thus, as expected from our theoretical arguments, the experimental results suggest that the assumption of independence of noise and message is not fully satisfied by neurons, *the noise depends on the message.*

3.5 Conclusions

It is believed that the dynamics of a real neuron can under certain conditions act as an integrator or under some other as a "coincidence detector" of fast input events. [Koch 98]. The first case corresponds to the conditions at the right side of Fig. 3.2 and the second to those at the left side. Thus a consequence of the results in Fig. 3.2 would be that the more "integrative" a neuron is, the more dependent on the signal the precision will be; the opposite is also truth: as long as these conditions are avoided temporal precision will be higher and independent of the signal.

A question that can profit from specific experimental work is whether or not the intrinsic dynamics of real neurons preclude the very existence of slow threshold crossings (or slower than some value) such that real neurons can always operate safely at the right side of the plot in Fig. 3.2 where uncertainty is independent of the message. It is very well known, for instance, that depolarization slower than a critical value does not originate action potentials, a property of excitable membranes known as "accommodation"[Mountcastle 74].

The relevance of the message-dependent nature of noise in spiking elements lies on its consequences for theories of brain function, and in particular for the usage of Shannon's Information Theory as a framework to measure the information content present in the output of a neuron. This is the subject of much work recently [Rieke 97]. In this regard, a basic assumption commonly made is that the noise introduced by the communication channel

is independent of the message being transmitted, which allows the modeling of a neuron as a Gaussian channel [Shannon 49]. This is indeed desirable given that departures from the Gaussian channel have proved to be rather cumbersome, and no general theory exists [Ash 90]. But what if we consider that this assumption is not entirely true? From our results it is safe to say that we cannot talk about the information content of one spike, because we have to know what is the context of the particular spike we are considering. This is particularly critical when the message consists of only a few spikes (for instance, in the auditory pathway), or in the event of fast conductance modulations [Borg-Graham 98], which can dramatically affect the reliability. Thus, perhaps the (un-suspected) most fundamental consequence of Mainen & Sejnowski's result [Mainen 95] is the implicit demonstration that cortical neurons are not always classical Shannon channels.

Part III

Of brains and languages

Chapter 4

Toward a Song Code

I call whatever represents some other object of thought a character. But that is said to represent which corresponds in such a way that from it something else can be thought, [and this] even if they are not similar, provided that all the things which occur in the one are referred, according to some definite rule or relation, to certain corresponding things in the other.

G.W. Leibniz, Philosophical Papers

By transforming the image from a space of simple discrete points to a congruent space where each equivalent point is described by the intersection of particular qualities in its neighborhood, we can then give the image in terms of distributions of combinations of those qualities. In short, every point is seen in definite contexts. The character of these contexts, generally built in, is the physiological synthetic a priori.

J.Y. Lettvin, H.R. Maturana, W.S. McCulloch and W.H. Pitts,
"What the Frog's Eye Tells the Frog's Brain"

Human language is, perhaps, the most fascinating outcome of the millions of years of evolution of the brain. Among the many viewpoints from which it can be studied, the most challenging one is its neurobiological implementation, under the assumption that all the external aspects of language

must have a material correlate. In this regard, there are three different problems to consider: the study of language itself, the choice of an experimental system, and the use of an appropriate analytic tool. These issues will form the backbone of the work presented in this chapter, focused on the study of auditory neural processing of song in birds.

4.1 Language and Animal Communication

A basic problem in the study of brain function is understanding the neural correlate of the external world. In other words, one wants to know the *code* that relates the functioning of the brain and the structure of the external world. In the case of language, the external world embraces most conspicuously two aspects. One, discussed in Chapter 1, is *syntax*, the set of rules that specifies how utterances are to be formed. The other, not totally unrelated, is *semantics*, which studies what utterances mean or are intended to mean. What is to be expected about the relationship between these linguistics structures and their neural correlate? We are well acquainted with the existence of maps in the brain: the topographic representation of the body in somatosensory cortices, retinotopic, ocular dominance and orientation maps in early visual processing, frequency and time-delay maps in auditory processing, etc. The picture that emerges is one in which the brain represents in an organized way simple features making up the complex objects present in the world. As for higher cognitive functions, i.e., the processing of com-

plex objects that constitute the meaningful units of the world (objects in the visual field, words in a sentence), not much is known from a physiological point of view. A common hypothesis suggests the duplication in higher stages of the strategy implemented in early processing: similar faces and other complex objects would map nearby in the Infero Temporal cortex following feature-based representational rules [Gross 92], as would semantically related words (establishing nested semantic classes, i.e. living objects, animals, four-legged animals, etc.) in different areas involved in language [Caramazza 96, Damasio 96]. Regarding these modes of cognitive processing, it is relevant to point that, despite its complexity, language has the advantages of the serial restriction on its output, and an extensive body of formalization of its structure and regularities, which after the 1950's and mainly under Chomsky's influence, ignited the hopes of a new synthesis between linguistics and physiology.

To what extent animal communication can be related to human language? This is certainly debatable, but no more than any other higher-order cognitive task (for instance, visual cognition) for which researchers turn to animal models in order to make inferences about the human brain. If one assumes that language and thought are inextricable woven together, that is to say, language is not a deus-ex-machina in the history of the brain nor an interface between cognition and the external world, one has to conclude that studying animal communication is also a way to understand the evolutionary origins

of thought. In this regard, there are strong similarities between pidgin languages (a contact language which is not the native language of its speakers), infant babbling, and utterances of trained animals like primates and parrots (for a good discussion, see [Bickerton 96]). This *protolanguage* appears to rest on the distinction actor-action-recipient, reminiscent of the grammatical construct subject-verb-object, but lacking any other fixed structure as well as the potentially infinite number of utterances which are characteristic of adult human language.

In spite of these concerns, animal communication systems pose in their own right several conceptual challenges from whose resolution much insight into language, as well as other cognitive functions of the brain, can be learned. The animal brain has to produce communication signals (be they auditory, visual, tactile, olfactory or a combination thereof) with reliability, it has to recognize incoming signals as such, separating them from 'noise' (background sound or smells, other elements of the visual scenery, etc.) as well as from signals produced by other species, and finally it has to extract the correct behavioral information that the signal is intended to convey (mating and warning calls are well known examples of conveyed behavioral information).

These challenges are illustrated by birdsong, a prominent case of animal communication. Song is a tool of key importance for the behavior of a large variety of songbird species. Among the various behavioral functions that song serves, we can mention: (a) adult song is unique to each individual,

acting as a signature that identifies the members of a colony; (b) the differential reproductive success of the males (i.e., the disposition of females to mate and the amount of eggs laid after mating) is determined by subtle properties of the song (some songs are more appealing than others); and (c) the broadcasting of song demarcates the limits of individual territorial domination [Immelman 69, Brown 75, Marler 82, Godard 91, Kroodsma 96]. It is evident then that the survival of the individual depends on a precise performance in the production and perception of song, and that both tasks imply complex brain functions.

One of the most salient features of birdsong is that its structure has a richness well beyond the distribution of frequency components, as appears to be the case in known primate calls [Rauschecker 95b]. The typical structure of birdsong consists of simple, recognizable units arranged in a score, making up a 'syntax', i.e. how the elements are related to each other, as well as a 'semantics', i.e., how the structure and the different behavioral or environmental situations are related to internal states of the animal other than auditory. Among the many species of songbird, canaries in particular have a highly structured song, consisting of a succession of phrases that are formed by the rapid repetition of basic components or syllables, reminiscent of those in human speech [Nottebohm 72, Nottebohm 78](Figure 4.1A). Several breeds of domesticated canaries exist, each with its own characteristic syllabic repertoire [Güttinger 85, Mundiger 95].

4.2 Snapshots of the elusive

Brain function is a collective phenomenon, and as such it should be (ideally) studied. In particular, the processing of complex stimuli like birdsong, must engage a large number of cells within and across brain areas. This is not to say that more reductionistic methods like single cell recordings should be abandoned; on the contrary, analytic and synthetic approaches must complement each other in a dialectic movement. Unfortunately, the techniques available to assess large-scale information processing in the brain are cumbersome and limited compared to the sheer complexity of the problem. The most promising among them seem to be functional magnetic resonance imaging (fMRI) and multi-electrode electrophysiological recordings, each providing a certain glimpse at brain function, together with inconveniences. In the case of fMRI, the advantages of non-intrusiveness and globality are compensated by a limited spaciotemporal resolution and by the fact that the signal is indirect, i.e., blood flow cannot be directly linked to the activity of individual cells. Moreover, restraint is unavoidable, which restricts the possibilities of studying fully behaving animals. This last restriction (together with the intrusiveness of the method) also applies to multi-electrode recording setups, compensated by the evident gain in quality and quantity of the collected information.

A different window into brain function is provided by the regulation of genomic activity in neurons. The gene family known as *immediate-early genes* (IEG) constitute the first wave of genomic expression upon changes in

extra-cellular signals, like the concentration of growth factors, and leading to changes in the pattern of expression in different cell types [Iyer 99]. In recent years it has also been shown that the electrical stimulation of neurons typically results in the sharp and transient induction of IEGs [Morgan 89]. Several of these IEGs encode transcription factors, i.e., genes whose protein products are localized to the nucleus, and whose function is the suppression or enhancement of the expression of other genes; moreover, the expression of IEGs has been correlated with a variety of plastic phenomena in neurons, such as changes in dendritic arborization (the sprouting of new dendrites) and long term potentiation (LTP, a long lasting change in the responsiveness of synapses) [Goelet 86]. Consequently, the expression of IEGs has been postulated to represent the early stages in a cascade of gene regulation leading to the modification of neuronal properties, which could provide a basis for memory formation.

Irrespective of its precise function, analysis of the IEG response has become increasingly popular as a tool in the assessment of sensorimotor responses to particular stimuli and of physiological states [Chaudhuri 97]. Some intriguing insights into neural mechanisms involved in vocal communication in birds have originated from the finding that the IEG *ZENK*, also known as zif-268 [Christy 88], egr-1 [Sukhatme 88], NGFI-A [Milbrandt 87] and Krox-24 [Lemaire 88], as well as the c-jun gene [Nishimura 88], are rapidly induced in the brain of songbirds when they hear playbacks of song

of the same species [Mello 92, Nastiuk 94]. This induction is most prominent in the caudomedial neostriatum (NCM) and adjacent caudomedial hyperstriatum ventrale (CMHV) [Mello 94], both areas thought to participate in auditory processing [Chew 95, Chew 96a, Chew 96b, Stripling 97, Vates 96]. *ZENK* induction has also been observed in several areas involved in motor control as a result of active singing behavior in captive as well as in wild songbirds [Jarvis 97b, Jarvis 97a, Jin 97].

As discussed in the previous section, one of the tasks the songbird brain has to solve is organizing the auditory stream impinging on the cochlea into a coherent neural representation, or song code. In the next sections we will show, using the expression of *ZENK* as a mapping tool to asses the state of the brain, that such a neural representation exists in NCM, and that its properties demonstrate the collective nature of song processing. The main finding is that information on the spatial distribution and staining levels of labeled cells is enough to provide for a classification of the expression patterns elicited by song syllables that **(a)** accords to the intra and interfamily relationships between the corresponding stimuli and, **(b)** confers salience to natural stimuli over artificial ones; moreover, **(c)** complex syllable maps cannot be reduced to any linear combinations of simple syllable maps. Finally, **(d)** these properties arise from the collective response of NCM's neurons to auditory stimuli, rather than from the behavior of single neurons in NCM. These results will allow us to conclude that this syllabic representation may

constitute a step towards deciphering the birdsong code [Ribeiro 98].

4.3 Syllabic Representation

The caudomedial neostriatum (NCM) is one of the main auditory processing stations in the songbird brain, located in the telencephalon, a region analogous, in evolutionary and functional terms, to the mammalian cerebral cortex. NCM receives inputs from the auditory thalamus and from field L, the primary auditory region of the telencephalon, and is indirectly connected with regions that originate a descending auditory pathway [Vates 96, Mello 98b]; therefore, it is likely analogous to supragranular layers of the mammalian cortex. Electrophysiological and 2-deoxyglucose uptake studies demonstrated responses to auditory stimuli, including song, in a vast zone of the avian caudal neostriatum, including NCM [Leppelsack 76, Theurich 84, Müller 85a, Müller 85b, Scheich 91]. Recent studies aimed more specifically at NCM show that the electrophysiological responses to novel songs habituate to repeated stimulus presentations in a song-specific and long-lasting manner [Chew 95, Stripling 97]. This habituation has been postulated as a mechanism involved in the formation of auditory memories [Chew 96a, Chew 96b]. NCM activation by song is also accompanied by a rapid and transient induction of the immediate early gene *ZENK* [Mello 92], also known as *NGFI-A* [Milbrandt 87], *zif-268* [Christy 88], *egr-1* [Sukhatme 88], and *Krox-24* [Lemaire 88]. *ZENK* induction in NCM is highest for same-species song,

suggesting a tuning to species-specific features. Furthermore, this genomic response is modulated by experience. It is highest when birds are presented for the first time with same-species but unfamiliar songs and habituates when the same song is repeatedly presented, to recur upon introduction of another novel song [Mello 95].

The *ZENK* gene encodes a zinc-finger transcription regulator. The corresponding protein has a structural motif that consists of a small number of amino acids grasping a zinc atom, and is responsible for the recognition of specific DNA sequences in the genome. As in most transcription factors, these sequences are typically less than 20 nucleotides long, providing specificity for the regulatory function. In neurons, the induction of *ZENK* occurs in association with electrophysiological activation, and has been used as a marker for brain activity in a variety of physiological and behavioral conditions [Chaudhuri 97]. However, *ZENK* induction is not synonymous with activity, but rather appears to reflect a specific type of activity related to neuronal and synaptic plasticity. As already mentioned, its induction is associated with long-term potentiation in the hippocampus [Cole 89] and neuronal morphological changes after exposure to complex environment [Wallace 95]. Indeed, electrophysiological activation in the songbird brain is not always accompanied by ZENK expression [Mello 94, Mello 95, Jarvis 97a, Mello 98a]. Nevertheless, no evidence is presently available for ZENK induction in mature neurons in the absence of electrophysiological activation (see Discussion

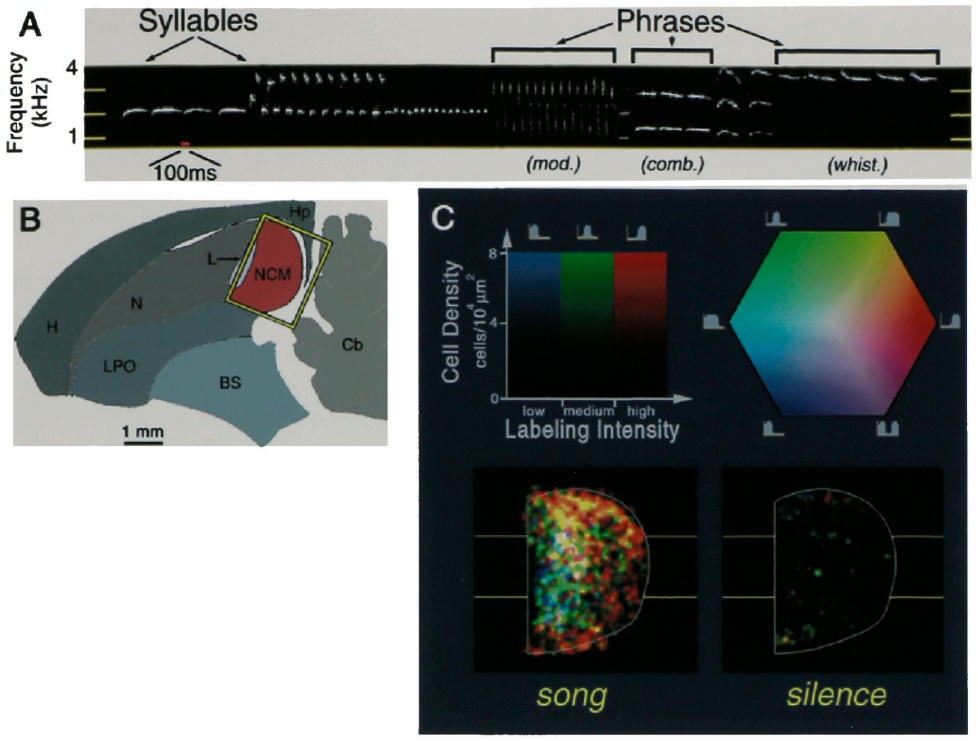

Figure 4.1: Canary Song Induces *ZENK* Protein Expression in the NCM of Canaries. (A) Sonogram (frequency vs. time plot) of the canary song studied here. The song consists of strings of phrases, which are themselves formed by the repetition of individual syllables that can be categorized in families. Examples of three syllable families (modulations, combinations, and whistles) are indicated. (B) Diagram of a parasagittal section of canary brain at the plane studied; the area enclosed by the rectangle indicates NCM as shown in all ZENK expression maps. Orientation: dorsal is up, and anterior is to the left. Abbreviations: BS, brainstem; Cb, cerebellum; H, hyperstriatum; Hp, hippocampus; L, subfield L2a of field L; LPO, lobus parolfactorius; N, neostriatum; NCM, caudomedial neostriatum. (C) Maps of ZENK expression in NCM evoked by song or after silence. Maps represent absolute numbers of labeled cell nuclei per unit area within three ranges of activation. The key on the upper left depicts the color assignment to the three populations of *ZENK*-labeled cells, and illustrates how brightness in the maps reflects cell density (in cell/10^4 μm^2; the key on the upper right indicates how mixed cell populations result in various color mixtures (hues). The small icons depict schematically the distributions of ZENK-labeling intensity that correspond to particular color mixtures. The NCM contour in the maps represents a standardized NCM outline.

in [Chaudhuri 97]). It is therefore reasonable to assume that ZENK reflects, to a large extent, the underlying activation state of the neuronal populations studied.

To investigate the auditory representation of song, we took advantage of the phenomenon of ZENK induction and mapped ZENK protein expression in the NCM of canaries presented with several song stimuli, including a whole song and several syllables that occur naturally in the repertoire of these birds. We studied the most frequent syllable types found in our aviaries; these cover three out of the four most common syllable types in this breed, according to a perceptual classification ([Güttinger 85]; see below). We grouped the syllables based on sonographic analysis into classes or families, namely whistles, combinations, and modulations (Figure 4.1A), and compared them with artificial stimuli. The resulting *ZENK* expression patterns were analyzed with a high resolution system that generated detailed and quantitative maps of the spatial distribution and labeling intensities of song-responsive neurons.

4.3.1 Whole Song

Presentation of whole songs results in marked and wide-spread *ZENK* induction in the NCM of songbirds [Mello 94, Mello 98a]. In agreement with those results, a large number of cells labeled for *ZENK* protein were seen in NCM in response to presentation of a whole canary song (Figure 4.1C), while much fewer cells, with an even distribution throughout NCM were observed in silence controls. The overall increase in the number of labeled cells in NCM

was ~7-fold (996 ± 130 labeled cells per parasagittal plane for whole song, compared to 140 ± 30 in silence controls; mean ± SEM). Our current mapping methodology has allowed for a much more detailed description of this pattern than reported previously. Practically all regions within NCM showed an increase in the density of *ZENK*-labeled cells, but this effect was not homogeneous: a particularly high density was observed rostrally, in the proximity of field L, whereas fewer cells were seen more caudally. Inhomogeneity was also observed in the range of labeling intensities: NCM's periphery showed a predominance of strongly stained neurons (red hues), while the central region of NCM was characterized by a more balanced proportion of cells high, medium and low labeling intensities (white hues), with some predominance of the medium range (green hues). As assessed by histograms of cell labeling intensity, the interanimal variability in *ZENK* staining patterns was rather low for whole song (SEM ~ 18 % of the mean labeling intensity; for details, see [Cecchi 98]) as well as for other stimuli tested (see Figure 4.5A below), which allowed us to compare the responses across different groups; very low interanimal variability was also observed for background labeling (SEM ~ 2.5 % of the means).

4.3.2 Syllables

Whistles

Individual whistles are naturally occurring narrow-band vocalizations at frequencies typically ranging between 1 and 4 kHz; these syllables belong to

classes 25 and 26 in Güttinger's classification [Güttinger 85]. We analyzed the expression patterns resulting from whistles of 1.4, 2.2, 2.8 and 3.5 kHz. In sharp contrast to the whole song, whistles resulted in activation that was restricted to a narrow domain within rostral NCM (Figure 4.2A, a-d). Each whistle activated a cluster of cells whose position varied systematically along the dorsoventral axis as a function of frequency: low frequencies mapped dorsally, high frequencies ventrally. These patterns were also characterized by a narrow range of labeling intensities, predominantly within the middle (green) bin.

Whistle Combinations

We included in this category canary song syllables composed of individual whistles combined either as sequences or chords (see Figure 4.2A, sonograms e, f and g, h respectively); the chords studied here belong to Güttinger's classes 25 and 26, while sequences are formed by two whistles of different frequencies (classes 25 and 26) occurring in quick succession. The maps resulting from both kinds of combinations were characterized by a band of activation that extended along the dorsoventral axis of rostral NCM (Figure 4.2A, e-h). Similar to whistles, combinations were also dominated by the middle (green) range of labeling intensities. Importantly, combining the same two whistles as a chord or as a sequence resulted in different patterns of activation (in Figure 4.2A, e$\neq$g, f$\neq$h); this difference was confirmed by a quantitative analysis of the spatial distribution of labeled cells (see below).

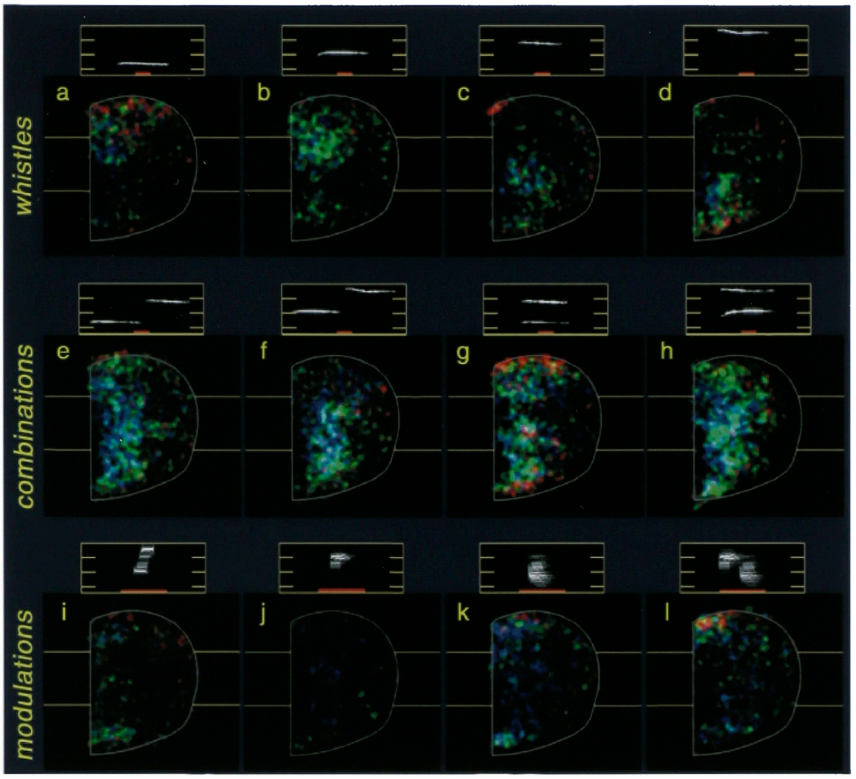

Figure 4.2: Syllabic auditory representation in the canary brain. Panels show maps of ZENK expression in NCM resulting from presentation of three families of song syllables: whistles (a-d), combinations (e-h) and modulations (i-l); the corresponding sonograms (frequency vs. time plots) are shown above each map. Each syllable family resulted in a family of ZENK patterns with distinct morphological properties (as detailed in the text). Notice that stimuli e and g are combinations of whistles a and c, while stimuli f and h are combinations of whistles b and d. Notice also that stimulus l (double sweep) is composed of sweeps j and k. Frequency scale in sonograms and color keys for maps are the same as in Figure 4.1; red bars in sonograms represent 100 ms.

Modulations

Modulations comprehend fast frequency sweeps that typically start between 2 and 4 kHz and can span up to 4 kHz within 100 ms, although they typically encompass $\sim$ 2 kHz in the songs recorded from our aviaries. The modulations we studied correspond to Güttinger's syllable classes 1 and 14. The resulting ZENK expression patterns were characterized by little overall activation of NCM. The majority of the observed labeled cells was in the low (blue) range of activation and tended to occupy the central portion of NCM; in contrast, cells with higher labeling intensities were mainly confined to NCM's periphery (Figure 4.2A, i-l). Although their frequency range encompassed the frequencies of the individual whistles studied here, modulations elicited little activation in the rostral NCM region where responses to the whistles mapped.

4.3.3 Natural vs. Artificial

The spatial distribution of the whistle maps suggested that rostral NCM could have a classical tonotopic organization. To test this hypothesis we compared the responses to whistles of frequencies 1.4 kHz and 2.2 kHz to those evoked by two kinds of artificial stimuli of corresponding frequencies: synthetic whistles and guitar notes. The dominant frequencies compared (1.4 and 2.2 kHz) were the same for the three stimulus types, but both kinds of artificial stimuli lack the small frequency modulations present in natural

whistles (Figure 4.3B). In addition, synthetic whistles have amplitude envelopes that resemble those of natural whistles but lack their slight amplitude modulations, as shown in Figure 4.3A. Synthetic whistles therefore represent an "impoverished" version of natural whistles, allowing us to compare the effect of the later with the response elicited by stimuli of identical dominant frequencies but fewer acoustic features. In contrast, guitar notes have the amplitude envelope characteristic of percussive sounds, i.e., a sharp increase followed by an exponential decay (Figure 4.3A). Sounds of this nature are not normally produced by the vocal apparatus of songbirds [Hartley 90, Goller 96] and are absent from the canary repertoire [Güttinger 85]; they therefore allow us to test for the effect of stimuli that are notably unfamiliar to the subjects.

As shown in Figure 4.4, the expression patterns elicited by artificial stimuli are substantially different from those resulting from natural whistles. Synthetic whistles and guitars resulted in a range of labeling intensities that was wider than that found for natural whistles. This effect can be best visualized in histograms of labeling intensities (Figure 4.5A); each stimulus type displaying a distinct cell distribution. The distributions of natural whistles are completely non-overlapping (Figure 4.5A, left), whereas the distribution of synthetic whistles is non-overlapping with those of either natural whistles or guitar notes for a large range og labeling intensities (Figure 4.5A, center and right). In this regard, synthetic whistles display an intermediate profile

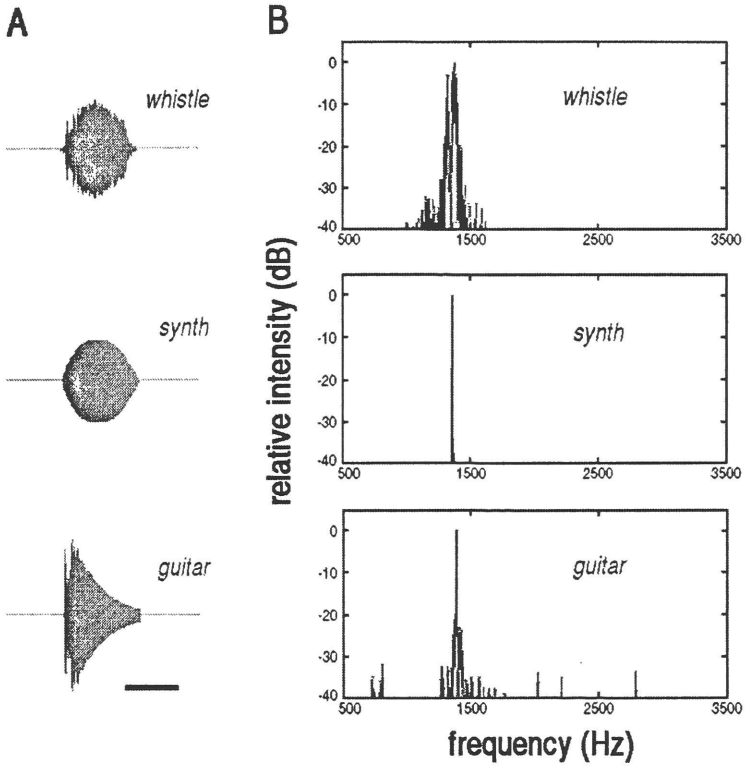

Figure 4.3: Comparison between Natural Whistle and Artificial Stimuli. (A) The amplitude envelopes of both natural and synthetic whistles have similar ellipsoidal shapes, but the slight amplitude variations typical of natural utterances are absent from the synthetic ones. The amplitude envelopes of guitar notes have a quick rise and an exponential decay. Scale bar represents 250 ms. (B) Fourier analysis of frequencies in natural whistles, synthetic whistles and guitars; a representative (1.4 kHz) of each stimulus type is shown. Contrary to the single frequency of synthetic whistles, natural whistles have slight frequency jittering around the dominant frequency. Guitar notes have a single dominant frequency; notice that their harmonic frequencies are 30 dB below the dominant frequency, as a result of filtering.

between natural whistles and guitar notes.

Patterns resulting from artificial stimuli were also not restricted to the spatial domains observed for natural whistles, but rather were gradually more spread (in Figure 4.4 follow the sequences a $\rightarrow$ m $\rightarrow$ o, and b $\rightarrow$ n $\rightarrow$ p): in contrast to whistle patterns, it was very difficult to discriminate the patterns resulting from artificial stimuli of different frequencies (1.4 kHz and 2.2 kHz). We quantified the topographic differences between patterns from natural and artificial stimuli using a clustering index (see definition of Appendix B). Figure 4.5B shows that the degree of clustering decreased and the number of labeled cells increased progressively as one moved from natural whistles to artificial ones, and then to guitar notes (notice the low interanimal variability within each group, as revealed by the small SEMs). We conclude that there is, in rostral NCM, a topographic organization of natural whistles according to their frequency. This organization, however, is progressively disrupted for synthetic whistles and guitar notes, providing a clear demonstration that frequency is not the only feature of a natural whistle being mapped in NCM.

4.3.4 Complex Stimuli and their Components

A quantitative analysis of combinations allowed us to investigate how the component parts of a complex syllable contribute to the representation of that syllable. Given that the component whistles occurred equally often during the presentation of combinations, one could imagine that combination maps were just the result of the summation of the patterns elicited by the

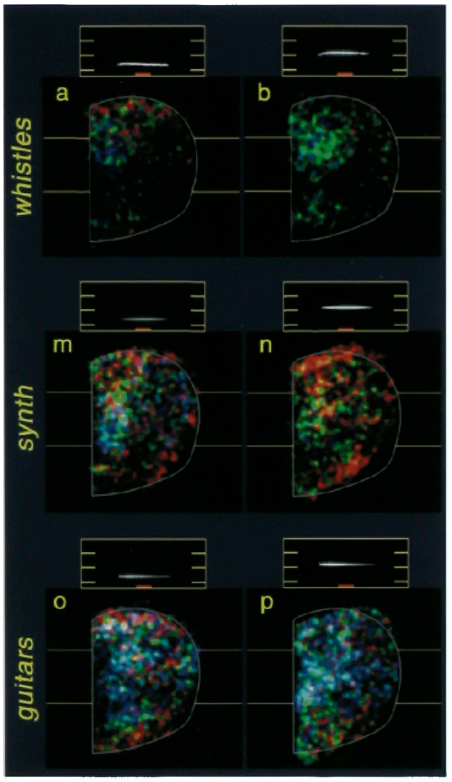

Figure 4.4: Tonotopy in NCM depends on natural features of the stimulus. *ZENK* expression maps resulting from natural whistles (a and b), synthetic whistles (m and n), and guitar notes (o and p) are compared. As one moves away from natural whistles towards artificial stimuli of corresponding frequencies, ZENK patterns become less clustered both in spatial distribution and range of labeling intensity. Maps a and b are the same as in Figure 4.2; color keys as in Figure 4.1.

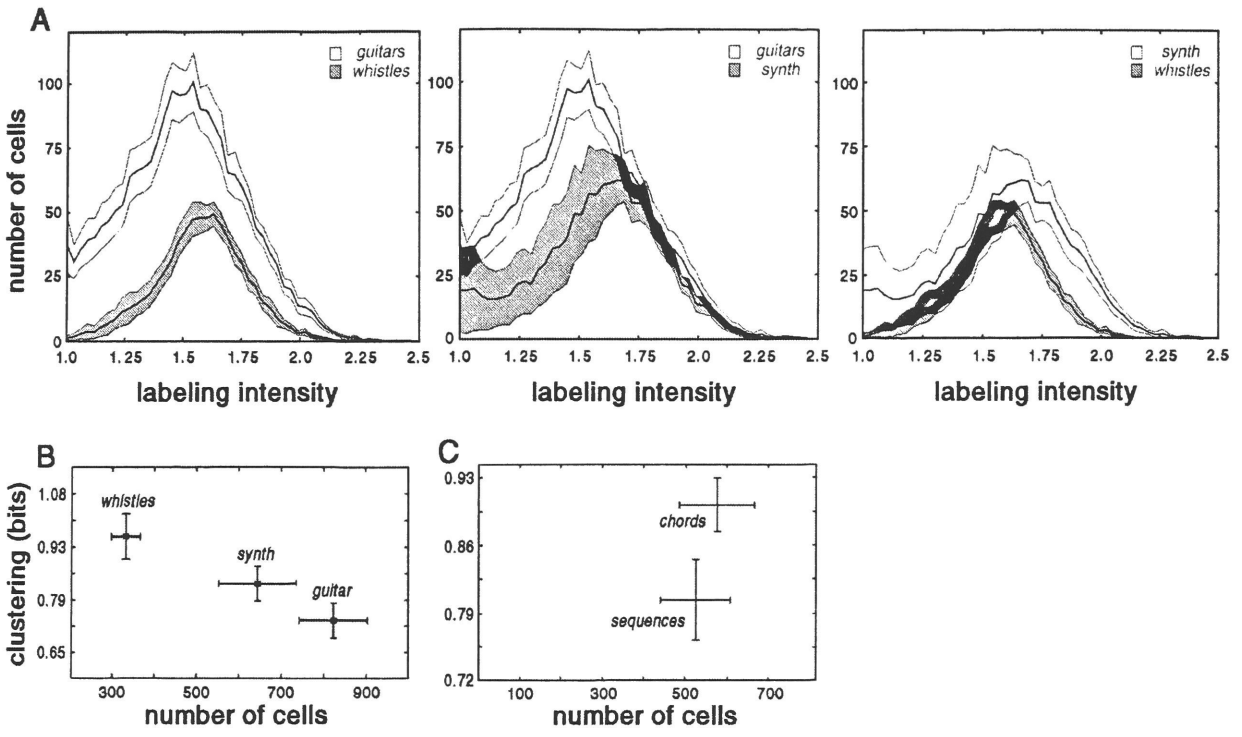

Figure 4.5: Quantitative Analysis of *ZENK* Expression Patterns. (A) Histograms of distributions of ZENK-labeled cells resulting from presentation of natural whistles, synthetic whistles and guitar notes. For each distribution a central line indicates mean values of the number of cells with a given labeling intensity, while shaded areas represent the variance (± SEM); regions of overlapping distributions in pairwise comparisons are indicated by the darker grey. Notice that each stimulus elicited a distinct profile: the distributions of natural whistles and guitar notes are completely non-overlapping ($p < 0.01$), while those of guitar notes and synthetic whistles are non-overlapping in the low to middle range of labeling intensities ($p < 0.05$), and those of synthetic and natural whistles do not overlap in the upper range ($p < 0.01$). (B) Clustering decreases and number of ZENK-labeled cells per section increases progressively from natural whistles to synthetic whistles and to guitars. Plotted are averages of the clustering indices of maps a and b (whistles), m and n (synth), and o and p (guitars) in Figure 4.4. Error bars represent SEM. (C) Sequences and stacks activate comparable numbers of cells in NCM, but can be discriminated by their degrees of clustering. Plotted are averages of the clustering indices of maps e and f (sequences), and g and h (stacks) in Figure 4.2. Error bars represent SEM.

component whistles. A comparison between a sequence map (Figure 4.6A) and the simple sum of the maps of its component whistles (Figure 4.6B) shows that this was not the case. The two maps have visible differences: for instance, the green area of activation in dorsal NCM (indicated by arrows in Figure 4.6) in the sum map is absent from the sequence map. When compared pixel-wise using a similarity index [Chapman 96] that scores 1.0 for total identity and 0.593 for two random patterns (see Experimental Procedures), these maps yielded a score of 0.79, providing quantitative evidence that the two maps are dissimilar.

It is possible, however, that the component whistles provided different contributions to the patterns elicited by their combinations. We therefore tested whether combination patterns could be the result of a weighted sum of the component whistles. We generated a map in which the contributions of the two component whistles were weighted so as to minimize the difference between their sum map and the sequence map (Figure 4.6C). This weighted map, however, still differs from the sequence map (compare Figure 4.6, A and C); its similarity to the sequence map (0.81) is just slightly higher than that of the simple map. This indicates that a combination is distinct from any linear superposition of the maps of its components. Further support for non-linear processing of combinations is given by the fact that the patterns elicited by chords or sequences of the same whistles are different, and can be clearly discriminated by their degree of spatial clustering (Figure 4.5C).

In addition, within the group of modulations, the map of a double sweep is different from the sum of the maps of its component sweeps (in Figure 4.2A, l≠j+k; for instance, the blue/green area of activation in the ventral-most portion of NCM in k is absent from l).

4.3.5 Habituation

To further address the relationship between a combination and its component whistles, we explored the phenomenon of habituation of the ZENK response after sustained stimulation [Mello 95, Mello 98a]. The rationale was to suppress the response caused by individual whistles from that seen when they are presented in combination, in order to reveal a combination-specific response. We exposed canaries to two whistles independently for a period sufficient to ensure habituation (7.0 hours), and then presented a combination (sequence) of the same whistles for 30 minutes; controls continued to hear the two whistles independently for the same period (Figure 7A).

While controls showed extremely low ZENK levels throughout most of NCM, canaries habituated to whistles and then presented with the combination showed robust ZENK activation, indicating a reinduction of gene expression by the last stimulus (Figure 7B). Interestingly, reinduction did not selectively reveal a subdomain of the activation in the sequence map. In fact, the most noticeable feature of the reinduction pattern is that labeled cells appeared mainly outside the region where the combination mapped prior to habituation (compare Figures 6A and 7B, lower panel). These results rein-

force the notion that a pattern resulting from a complex stimulus cannot be trivially explained by the patterns elicited by its components.

4.3.6 Principal Component Analysis (PCA)

In order to obtain an independent quantitative confirmation of the overall representation of song and syllables in NCM, the ZENK expression patterns were subjected to Principal Components Analysis (PCA). PCA identifies "features" (components) that best discriminate patterns within a given set [Richmond 87], essentially by performing a high-dimensional version of the linear best fit. The data are represented by an ensemble of n-dimensional vectors, $\{\vec{x}\}$ (in our case, the components refer to the number of cells and level of activation across the topography of the map, see Experimental Procedures). One can always express the vectors in an orthogonal basis:

$$\vec{x} = \sum_{i=1}^{n} c_i \vec{\phi}_i$$

where $\Phi = \{\vec{\phi}_i,\ i = 1, n\}$ are n-dimensional, $\vec{\phi}_m \cdot \vec{\phi}_n = \delta_{mn}$, and $\{c_i,\ i = 1, n\}$ are the components of $\vec{x}$ in the basis Φ. Given the orthogonality condition, the components can be computed as $c_i = \vec{x} \cdot \vec{\phi}_i$. The question that one wants to answer is the following: how well can we reconstruct $\vec{x}$ if only $m(< n)$ components are computed? In other words, we want to minimize the square error ϵ between $\vec{x}$ and the reconstructed $\vec{x}^*$,

$$\vec{x}^* = \sum_{i=1}^{m} c_i \vec{\phi}_i + \sum_{i=m+1}^{n} b_i \vec{\phi}_i$$

$$\epsilon^2 = \langle (\vec{x} - \vec{x^*})^2 \rangle = \langle \sum_{i=m+1}^{n} (c_i - b_i)^2 \rangle$$

where $\langle \cdot \rangle$ means average over the data ensemble. The optimal choice for the basis Φ is the one formed by the eigenfunctions of the correlation matrix,

$$\mathcal{C} = \langle (\vec{x} - \langle \vec{x} \rangle) \cdot (\vec{x} - \langle \vec{x} \rangle)^T \rangle$$

$$\mathcal{C}\vec{\phi}_i = \lambda_i \vec{\phi}_i$$

while the optimal choice for the b_i is given by $b_i = \phi_i \cdot \langle \vec{x} \rangle$. Finally, the reconstruction error is given simply by

$$\epsilon^2 = \sum_{i=m+1}^{n} \lambda_i$$

The expansion then produces a natural classification of the components $\{\phi_1, \phi_2, ...\}$ according to the value of the eigenvalues $\{\lambda_1 \geq \lambda_2, ..\}$, so that the first components are the 'principal' ones.

Figure 8 depicts two different perspective projections onto a plane of the 3D space defined by the first three components of the PCA of the expression maps. The first component obtained (x axis) was related to the number of cells, the second (y-axis) to intensity distribution and rostrocaudal activation, and the third (z-axis) to dorsoventral activation. We found that these first three components were already enough to classify and organize the patterns into a structure fully concordant with that of the set of stimuli: different

families of patterns lie at different depths/strata, without any disruptions of one family of patterns by other families. For instance, the family of patterns obtained from whistles spans a thin tube along the z-axis that does not intersect the subspaces defined by other families. Of particular interest is the fact that PCA sharply distinguishes whistle patterns from those of similar but artificial stimuli (synthetic whistles and guitar notes). In addition, the PCA arranged individual members orderly within their families, preserving stimuli kinship. Notice in this regard how the third component organized whistles and combinations according to their frequency, and the modulations according to their ending frequency. For artificial stimuli, synthetic whistles were shifted in the z-axis respect to whistles of same frequencies, but still aligned along this axis according to frequency, whereas this property was disrupted for guitars.

4.4 Properties of the Representation

4.4.1 NCM is Not a Classical Tonotopic Map

Different canary whistles activate clusters of *ZENK*-labeled neurons in rostral NCM; these clusters map along the dorso-ventral axis in a graded manner, according to their frequencies. This finding fits well with the fact that rostral NCM receives afferents from the adjacent primary auditory telencephalic region, field L, which is tonotopically organized [Müller 85a, Heil 91]. It would thus seem that rostral NCM follows a classical tonotopic organization

[Woolsey 42, Reale 80], with one topographic axis directly mapping sound frequency. If this were the case, different stimuli of same frequencies should have similar ZENK patterns. The comparison between natural whistles and artificial stimuli, however, argues against a simple tonotopic hypothesis. Synthetic whistles and guitar notes elicited widespread activation in NCM, in contrast to the highly clustered patterns obtained with natural whistles. Indeed, the degrees of clustering for natural whistles, synthetic whistles and guitar notes are significantly different. Furthermore, the broadening of the spatial distribution of *ZENK*-labeled neurons was accompanied by a corresponding broadening of the labeling intensity distribution: while whistle patterns are dominated by intermediate (green) levels of activation, artificial stimuli patterns have considerable amounts of cells with low (blue) and high (red) *ZENK* levels.

The patterns resulting from artificial stimuli are not just stronger versions of the patterns resulting from natural whistles of corresponding frequencies. If this were the case, one would have expected obvious areas of peak activation in the former, corresponding to the clusters seen in the whistle maps. A close inspection of Figure 4, however, shows that guitars had no readily recognizable areas of peak activation; as for synthetic whistles, besides a considerable degree of activation in caudal NCM, most of the activity in rostral NCM is not clustered in the same position as in the corresponding natural whistle patterns, but is rather shifted. This effect was quantitatively

confirmed by PCA: the position along the axis that correlates with frequency is different for natural and artificial stimuli of same frequencies (compare the position along the z axis of patterns a, m and o, and b, n and p in Figure 8).

In summary, the specificity of the ZENK response is degraded as one moves away from the natural stimulus: while it is trivial to discriminate two whistles of different frequencies based on their ZENK patterns, this task is more difficult for synthetic whistles of same frequencies, and virtually impossible for guitar notes. These observations constitute the basis for our claim that NCM does not behave as a classical tonotopic map: its topographic organization according to frequency is dominated by natural features of the stimuli. Interestingly, electrophysiological studies of field L neurons in the mynah bird [Hose 87] have demonstrated a tuning to pitches and rhythms typical of animal communication sounds. It is possible that similar mechanisms operate in NCM, but further work is necessary to establish the precise acoustic features to which NCM neurons are tuned.

4.4.2 Clustering

The clustering index provides a useful readout of the topographic relationships among labeled cells in the *ZENK* patterns. It is independent of the number of cells in the patterns [Okabe 92], reflecting only their relative spatial distribution. For instance, two patterns with comparable number of cells may have substantially different clustering indices (see Figure 5C) and vice versa. Thus, the inverse relationship between clustering and number of cells

seen in the comparison of natural and artificial stimuli (Figure 5B) does not reflect an intrinsic interdependence between these two measures, but rather the progressive change in the spatial distribution of labeled cells.

This index corresponds to the information content of the spatial pattern of labeled cells ([Shannon 48, Brillouin 89]; see Appendix). An intriguing conclusion that can therefore be drawn from our analysis is that the expression patterns elicited by natural syllables contain more information, or salience, and that conversely those elicited by artificial stimuli have less information or salience, being hardly distinguishable from one another. Information Content cannot be equated with biological relevance or meaning, and it should be clear that we do not imply that patterns resulting from natural stimuli are more meaningful or relevant in any behavioral sense. This could, in principle, be said of the stimuli themselves, but not necessarily of their corresponding *ZENK* expression patterns. The higher Information Content refers solely to the internal structure of the ZENK patterns, which in turn presumably reflects the functional organization of NCM. The higher clustering in the patterns elicited by natural whistles, as opposed to artificial ones, provides a topographic parallel to results showing that spike trains elicited by naturalistic stimuli carry high Information Content in their temporal structure (see Chapter 3). Obviously, we cannot assert that the mechanism described in Chapter 5 is responsible for this phenomenon; rather, we would like to suggest that clustering, both in the temporal and spatial domains, may be a

relevant variable to consider when investigating brain representations [1] which could eventually be functionally linked [2]. We will return to this issue in the next chapter, within the context of a proposed model of neural organization that accounts for some of the phenomena described.

4.4.3 Non-Linearity of Complex Stimuli

A recurring theme in our study was the systematic non-linearity observed whenever we broke a stimulus into smaller (or simpler) pieces: the maps of natural whistles cannot be reduced to the maps of synthetic whistles, nor maps of combinations to those of individual whistles, or maps of double sweeps to those of individual sweeps. Hence, a complex syllable pattern is related to, but cannot be fully explained by the patterns of its components, indicating a non-linear processing of auditory information. It is tempting to speculate that non-linear processing of discrete syllable features provides a mechanism for building up neural responses with unique identity within NCM. It is at present unclear whether such processing occurs somewhere along the ascending auditory pathway, within NCM itself, or as a result of interactions between NCM and its reciprocal targets.

The habituation experiment provided strong support for the notion of

[1] There is a growing literature on theoretical aspects of code sparseness for which our data on clustering could be of relevance. See, for instance, [Olshausen 96, Földiák 95]). For an application of knowledge on neuroarchitectonical constraints to brain modeling, see [Durbin 90].

[2] One possible connection, discussed in Chapter 3, is through the patterns of calcium signaling needed to activate transcription factors.

non-linearity. It was conducted as an independent and direct assessment of whether a pattern elicited by a complex syllable can be explained by the responses evoked by its component parts. The results show clearly that this is not the case: one cannot abolish the response to a sequence by suppressing the responses to its component whistles. We therefore conclude that combinations are processed as distinct syllables, with unique identity. The fact that a sequence presented after habituation to its component whistles activated mostly NCM regions that were not activated by the same sequence when presented without previous habituation indicates that the functional architecture of NCM can be modified by experience or context. A possible explanation for such reorganization, given the existence of GABAergic networks in the avian telencephalon [Müller 88], is that presentation of individual whistles could have an inhibitory action on caudal NCM regions. According to this hypothesis, habituation of rostral NCM by repeated presentation of individual whistles would release the inhibition on caudal NCM, leading to the *ZENK* expression pattern observed in Figure 7B (bottom). Although further experimentation is necessary to elucidate the mechanisms involved, our data clearly indicate that syllabic maps are not the result of a static parceling of NCM.

Another noticeable feature of the ZENK expression patterns elicited by different stimuli was that they partially overlapped. In consequence, particular NCM regions must participate in the representation of more than one

syllable. Likewise, it is probable that single neurons take part in the response to more than one stimulus. While the issue awaits confirmation, indirect evidence for the participation of single NCM neurons in the response to multiple song stimuli has been presented elsewhere [Mello 95].

The properties of non-linearity and partial overlap suggest that syllabic representations in NCM involve sets of highly interacting cells, recruited from the total pool of NCM neurons; each stimulus would recruit a subset of the total neuronal population in a unique manner. According to this hypothesis, the response properties of individual cells cannot be meaningfully described without reference to the overall physiological state of the brain region in which they occur. Indeed, it should be emphasized here that non-linearity and partial overlaps were properties we observed while analyzing the global patterns of activation, and not the responses of individual cells that participate in them. Thus, these properties emerge from the collective behavior of neurons within NCM. Such a scheme is in line with the notion of ensemble coding [Lashley 50, Gross 92, Nicolelis 95, Deadwyler 97, Joerges 97, Rolls et al. 1997]. It also resonates with recent observed non-linear combinatorial interactions of odour responses in the honeybee brain, postulated as possibly "crucial for the formation of singular codes for complex odour blends" [Joerges 97]. In our study, we provide strong and quantitative evidence in favor of such a coding scheme in the brain of a higher vertebrate.

4.5 Gains, Loses and Caveats

4.5.1 Gene Mapping vs. Electrophysiology

The *ZENK* mapping methodology allows one to assess the activation levels and topographic distribution of large populations of song-responsive neurons. Particularly noteworthy is the fact that the experiments can be done in alert animals, without interference with their natural behaviors. Furthermore, *ZENK* levels in silence or after repeated stimulation are very low, yielding a high contrast for detecting responses to novel stimuli. In the present study, an average of 450 labeled neurons were detected per parasaggital plane, embedded in a lattice of $\sim$ 6000 nonlabeled neurons. The main disadvantages of the method are the poor temporal resolution and the fact that one bird is killed for each stimulus, limiting the number of stimuli that can be included in one study. In contrast, single unit electrophysiological studies can asses the response to several different stimuli and provide detailed information in the temporal domain but typically use anesthetized animals and investigate the activity of < 40 neurons per animal (e.g., [Margoliash 92, Solis 97, Doupe 97]; see also discussions in [Capsius 96, Margoliash 97]). The information obtained with the ZENK mapping methodology is hence complementary to that obtained with electrophysiology. This tradeoff proved invaluable for our purposes, as the ability to recognize and classify *ZENK* patterns resulting from different syllables strictly depended on analysis of the collective behavior of a large population of syllable-responsive neurons.

As noted in the introduction, ZENK expression cannot be equated with electrophysiological activity. While immunostained regions in NCM most likely represent neuronal activation, areas devoid of ZENK could, in principle, have undergone neuronal depolarization without triggering the ZENK response. As has been suggested for other systems [Dragunow 96], the ZENK patterns we described rather may reflect a particular type of neuronal activity associated with plasticity, perhaps leading to the formation of auditory memories [Chew 95, Mello 95], but further work is needed to settle the issue. Indeed, the occurrence of *ZENK* induction associated with both learned and unlearned components of sexual behavior in the Japanese quail [Ball 97] indicates that caution needs to be exercised when trying to establish links between *ZENK* gene expression and learning. The interpretation of our present results, however, does not depend on determining the precise relationship between *ZENK* induction and electrophysiological activity or neuronal plasticity. In fact, as discussed below, a quantitative analysis of the intrinsic organization of the patterns per se can provide some important insights into the functional organization of NCM.

4.5.2 Behavioral Responses

Behavioral studies have demonstrated that female canaries show sexual preference for own-breed songs, scored as the number of copulatory solicitation displays evoked by songs of various breeds, as well as other species' songs [Vallet 92]. Female canaries also show sexual preferences for particular

phrases within a song repertoire, mainly responding to short phrases with abrupt frequency modulations and short silences [Vallet 95]. We did not observe that fast modulations elicited stronger responses than other syllables assessed; in fact, modulations yielded the weakest *ZENK* responses in NCM. Thus, even though NCM appears to play a role in the auditory processing of syllables, other brain sites may show a more selective response to song elements with high sexual relevance. It is important to note, nonetheless, that the females studied here were not in the state of sexual responsiveness to males during which the preference for modulations has been described [Vallet 95], and which is characterized by behaviors such as nest building and egg laying and by particular hormonal changes [Goldsmith 84, Wingfield 90].

4.6 Toward a song code

The *ZENK* pattern elicited by a whole canary song (as in Figure 1C) is very complex both in spatial distribution of labeled cells and in the range of labeling intensities. The pattern presumably reflects the highly structured nature of song itself, but one cannot decode such a relationship based on that results alone. We took in our study a reductionistic route and dissected the song into basic components. The results obtained provide a possible substrate or support for a syllabic representation in NCM. Individual ZENK maps are representations of particular syllables, and could be considered as the outputs of a syllabic code, in its turn understood as the rules by which

the brain transforms the physical properties of a set of syllables into a set of representations. By analogy, one could say that a single syllabic map is like a letter in the alphabet, but not the alphabet, much less the language. Indeed, we have not conducted an exhaustive survey of the entire canary repertoire, nor do we claim to have uncovered all of the rules that govern the transformation of song syllables into their representation. We have rather deduced, from the intrinsic organization of syllabic maps, some properties of a syllabic code.

The results obtained from the PCA constitute a central part of our argument. As it was fed only information on the spatial distribution and labeling intensities of ZENK-labeled cells in NCM, the PCA provided an independent and unbiased confirmation that sufficient information is present in ZENK expression patterns to discriminate the auditory responses to several natural canary syllables, preserving their intra- and interfamily relationships. Thus, ZENK patterns in NCM not only display syllabic identity, but also a higher order kinship organization, based on families of similar stimuli. The ability to perform such a classification task is a minimum requirement for an auditory brain region where a syllabic representation takes place, and is met quite nicely by NCM.

It should pointed out that the classification scheme we adopted for syllables is based on sonogram analysis, without any assumptions about the correspondence to perceptual categories. The purpose here was strictly an-

alytical: to learn how the representation of a complex syllable is related to those of its component elements, it was important to separate syllables composed of multiple whistles (which we named combinations) from whistles that occur naturally without accompanying elements (named whistles). Our classification segregates chords from whistles of the same fundamental frequency, even though such syllables are placed in the same category by a classification based on a combination of sonogram analysis and human perception of canary songs [Güttinger 85]. Conversely, our classification places together, within a single family, syllables such as sequences and chords, which are clearly perceived as distinct by humans. Our results demonstrate that the organization of syllabic maps in NCM mainly accords to such sonographic classification, indicating that NCM is fundamentally responsive to spectral features of the stimuli. Interestingly, the patterns elicited by chords and sequences made of the same components can be distinguished by their degree of clustering. In that respect, chords are closer to individual whistles than to sequences (Figures 5B and 5C), in agreement with the fact that humans perceive chords as less similar to sequences than to whistles of the same fundamental frequency. Further experimentation will be necessary to elucidate the relationship between *ZENK* patterns in NCM and the bird's ability to perceive and discriminate syllables.

How do maps elicited by individual syllables relate to the complex pattern elicited by a whole song? Can different songs also be discriminated based on

ZENK expression patterns in NCM? Is that information then made available to brain centers more directly involved in the bird's behavioral response to song? Our study has uncovered basic features of NCM's functional organization, and provided tools that allow such questions to be addressed, taking us a step closer to deciphering the brain encoding of birdsong.

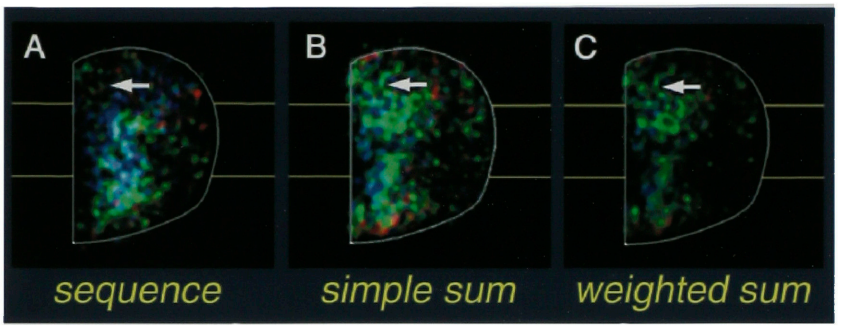

Figure 4.6: Contribution of the maps of two whistles to the ZENK expression map of their combination as a sequence. The map resulting from the presentation of a sequence (A; same as f in Figure 4.2) is compared to the simple sum (B) and weighted sum (C) of the maps of the corresponding component whistles (b and d in Figure 4.2). Notice that neither sum map is a good approximation to the sequence map; in particular, maps (B) and (C) contain a region of green activation in dorsal NCM (indicated by the arrows) that is not present in (A).

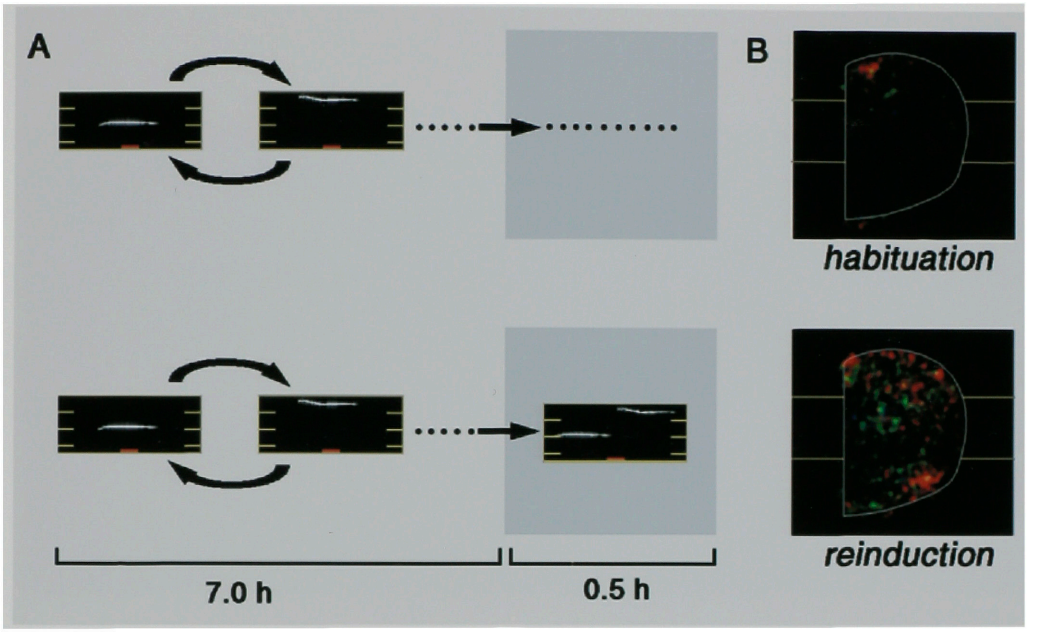

Figure 4.7: Habituation Experiment. (A) Experimental design: canaries were presented for 7.0 h with two component whistles (same as b and d in Figure 2), alternated every 30 seconds; during an additional 0.5 h period, controls continued to hear the whistles in alternation (upper row), while experimental animals were presented with the same whistles combined as a sequence (bottom row); all animals were killed 60 minutes after the offset of stimulation. (B) While ZENK expression is nearly suppressed in NCM after habituation to the component whistles (upper panel), presentation of the sequence after habituation to the individual whistles causes ZENK reinduction (lower panel); the latter is most pronounced in NCM regions which were not activated by presentation of the sequence without previous habituation to its component whistles (compare with Figure 6A). For both panels, color keys as in Figure 1C.

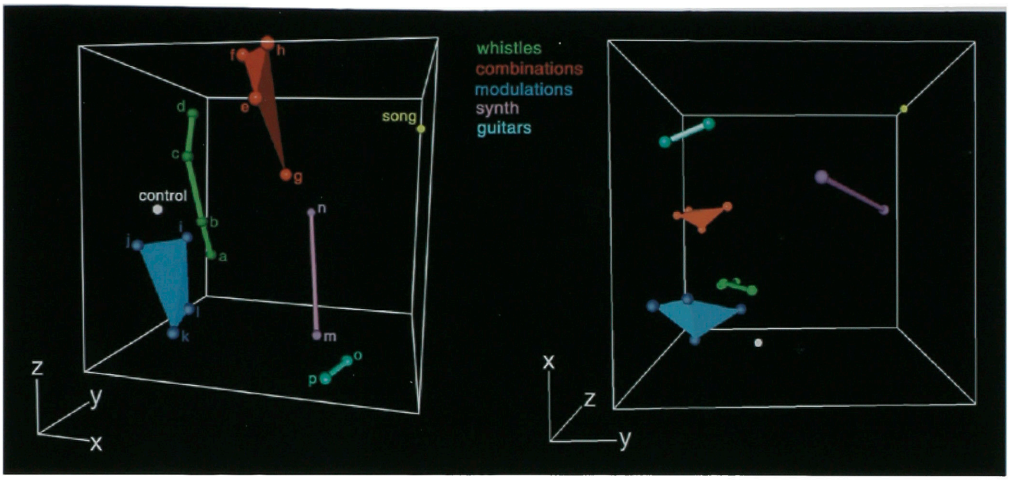

Figure 4.8: Principal Component Analysis (PCA) of *ZENK* Expression Patterns Resulting from Natural and Artificial Stimuli. Shown are two perspectives of the space defined by the first three components of PCA (represented by the x-, y- and z-axis respectively); the components were normalized independently so that the coordinates span a unit cube. The distribution of ZENK patterns within this 3-D space mirrors the organization of the stimuli into different families; notice that individual patterns are aligned along the z-axis according to their frequency (natural and synthetic whistles) or ending frequency (modulations). The letters correspond to patterns shown in Figures 2 and 4.

Chapter 5

Coding or the language of neural space

In the beginning was the Logos, without the Logos was nothing made that was made...

St. Augustine

Half a proposition does not stand to a proposition as half a bread-roll to a bread-roll but as 'half a knight's move' to a knight's move, and there is no such thing as 'half a knight's move'.

L. Wittgenstein, Philosophical Grammar

The anatomic variability is so great as to preclude, I believe, any theory which assumes regularity and precision of anatomic arrangement. Such facts lead me to believe that theories of neuron interaction must be couched, not in terms of the activity of individual cells but in terms of mass relations among cells. Even the simplest bit of behavior requires the integrated action of milliones of neurons; the activity of any single neuron can have little influence on the whole, just as the path of an individual molecule of gas has little influence on the gas pressure... We shall probably have to use a different kind of model, a model which can be explained in principle by individual neuron action but which involves a somewhat different set of concepts and laws of action.

K. Lashley

The discussion presented in this chapter stems from the need to clarify the concept of *neural code*, and to further explore the consequences for issues of coding of the results presented in the previous chapter. It is evident that the concept of code has multiple and conflicting interpretations in the neuroscience community. The confusion about its meaning taints the field, and is to a large extent caused by the complexity of the problem and the lack of common experimental and theoretical frameworks. In the remainder of the chapter we will discuss different approaches to issues of code, ensemble and single-cell coding, and self-organization of neural space, confronting them with the results on song representation.

5.1 Different definitions of code

Code has been used very loosely in the neurobiological literature, its meaning varying according to the specific context. Particularly revealing of this state of matters is the frequent confusion between *code* as an *encoding algorithm* and *code* as the actual *encoded information*. An example of the latter usage is [Konishi 91]:

> *"Neural codes are pieces of information that neurons convey to other neurons."*

Even when code is used as the *encoding algorithm*, it may assume different

meanings: first, code as the *spike language* used by neurons to communicate, as illustrated below [Rieke 97]:

> *"We have defined the problem of neural coding in terms of one basic question: given the spike train $\{t_i\}$, what can be said about the unknown stimulus waveform $s(t)$?".*

Second, code as the set of rules that specify how external information is mapped by the neural space: [Rauschecker 95a]:

> *"In order for a behavioral reaction (...) to be appropriate, it is necessary that (...) the same code is used regardless of sensory modality. Thus, a cortical module at any one processing level applies the same type of operation to different types of input and transforms them into a specific response.".*

What is then a brain code? One should find a precise and at the same time embracing definition, likely to clarify new concepts and guide the research. Along this line, the question to ask is what is the problem the brain has to solve for which a code is necessary. The task of the brain is not only to gather and transmit information, but also to consistently organize it, at the same time discarding redundant and unnecessary pieces. A brain code consists not only of a one-to-one mapping of the external world into neural space: the entire structure of the (relevant) world is also represented, and that includes relationships between different elements. In the case of birdsong, it

should consist of the rules of transformation of a set of stimuli into a set of representations. In other words, a brain code defines a topology of the neural space, whereby each element (represented) has a position relative to the rest. This basic feature would allow the brain then not only to organize the "known world", but also to map new pieces of information impinging on it. This definition approaches the meaning given above in [Rauschecker 95a], and refers to the rules of transformation used to represent song syllables and other stimuli in NCM.

What is the input space? The brain receives sensory input from the external world through different peripheral receptors: photoreceptors for vision, mechanoreceptors for touch, olfactory receptors for olfaction, hair cells for audition. The actual input to the central nervous system is the output of these sensory systems. That is to say, what the CNS knows about the auditory world comes to it through the information provided by the cochlea, which "translates" sound waves into electric signals. The cochlea dictates the topology of the auditory input, so that pure tones at 440 Hz and 460 Hz are more similar than one at 600 Hz, and so on. It is then the task of the brain to take this input and "interpret" it as auditory information. To gain insight into the topology of the input space corresponding to song syllables, one can model the cochlea as a linear Fourier analyzer, providing the auditory pathway with an instantaneous frequency spectrum of the incoming sound, integrated over a time window of 20 $msec$ [Morgan 91]. An approximation to

the input space topology is given by the principal component analysis of the set of syllables used in our study; that is to say, we cast the ensemble of the synthetic cochlea into a linear space. The results of preserving the components corresponding to the first three eigenvectors are shown in Figure 5.1. It can be appreciated how, in cochlear terms, natural and artificial stimuli are very similar. This can be understood as the differences between, for instance, a natural whistle and the synthetic one (raggedness of the envelope) are very small when compared with differences across stimuli (other syllables in this case). Moreover, the assumption of linearity leads to a cochlear representation of combinations as interpolations between the representation of the components, which can be seen by their position in the PCA space, as opposed to the allocation into unique subspaces revealed by the analysis of expression patters discussed in the previous chapter. Altogether, the comparison between the synthetic cochlea and the expression patterns suggests that an active reorganization of the input space grants salience to the syllabic identity of the stimulus.

The purpose of this exercise, though, cannot be but didactical. On the one hand, even though the Fourier approximation is a common one, we lack precise information on the cochlear output, in particular the details of the temporal encoding of sounds. On the other hand, one should know the entire statistical ensemble of, at least, all the songs to which birds in a particular colony are exposed to, and which in principle they should discriminate.

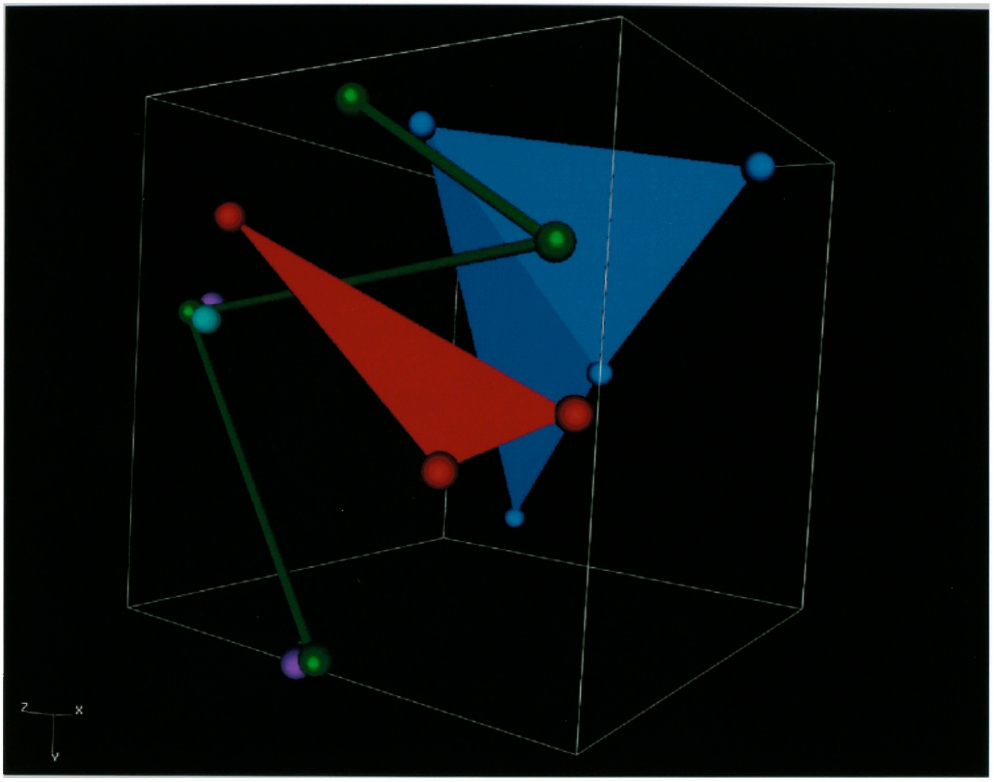

Figure 5.1: First three components of Principal Components Analysis of the sonograms of all canary syllables studied in the previous chapter. Green corresponds to natural whistles, purple to tone, light blue to guitar tone, red to combinations, and blue to modulations

There are two reasons for this, not altogether independent, and indebted to Aristotle's metaphor of the mind as a clay *tabula rasa* on which the world carves *impressions*, molding it, as opposed to the platonic realism of the ideas, which have an existence outside and independent of the mind. The first reason is the hypothesis of *optimal processing*: according to this idea, the brain adapts its response to the statistical properties of the input so as to use its limited resources to maximize the efficiency of signal processing. In [Barlow 61] it was proposed that this is the design principle of retinal ganglion cells: fibers in the optic nerve must use their dynamic range to carry as much information of the visual field as possible. Nearby points of the retina are statistically correlated and therefore have redundant information; the design of the ganglion cells is such that redundancy is minimized (which amounts to the statistical decorrelation of the fibers), with the constraint of keeping maximal information of the visual field [Atick 92]. Applied to songbirds, this idea suggests that the auditory pathway is optimized to transmit and process song, according to the statistical properties of the song ensemble.

The second reason is the hypothesis of *self-organization* of the brain. The rationale is that it would be impossible to store in the genetic material the precise "schematics" to layout the structures found in the brain. Therefore, the only sensible alternative is an algorithm based on a few (genetically determined) rules acting as a scaffold for the organization of the neural space according to the properties of the input. This is the idea behind most of the

investigations on ontogenesis and plasticity of neural responses, and on self-organized artificial neural networks. To illustrate the idea, we will present one of the most appealing (for its simple physiological implementation) models of self-organization, the Kohonen map [Kohonen 96]. It consists of a feature layer of model neurons receiving input from an input layer; upon presentation of an "example", the unit with the best response in the feature layer is chosen as the winner, and its synaptic weights rewarded on a Hebbian basis (coincidence of pre- and post-synaptic activity). The immediate neighbors of the winner are also rewarded, but in amounts inversely proportional to the distance to the winner. Successive presentations of examples drawn from the input ensemble result in the formation of a map in the feature layer (equivalently, neural space) that: **(a)** preserves the topology of the input in neural space (examples with similar features are neighbors), and **(b)** tessellates the neural space in feature-dominions whose size is proportional to the density of the distribution of input examples. These two properties approach the requirements for a neural code previously stated. Moreover, the physiological implementation of this mechanism is extremely simple, requiring, besides the Hebbian learning, short-range lateral excitatory (to implement the continuity of the neural map) and long-range inhibitory (to implement the winner-take-all strategy) connections, a common motif also known as *Mexican hat*. The existence of feature maps in different areas of early processing (most prominently, in the somatosensory cortex) sharing the aforementioned properties

(a) and (b) seems to support the self-organization hypothesis. Returning to the case of songbirds, this hypothesis suggests that the brain is shaped by the auditory stimulation, including song, that particular individuals were exposed to. Therefore, based on a detailed knowledge of the input space, one should be able to predict to a large extent the structures found in the songbird brain. Preliminary work [Ribeiro 98] suggests that maps in NCM are strongly influenced by the properties of the input, both during maturation and in adulthood. We will return to this in the last section of this chapter.

Finally, regarding the neural representation, we also lack temporal information, as well as information on areas other than NCM, which may be necessary to establish a complete auditory representation (we know that other areas are necessary for the processing). Nevertheless, we hope that this preliminary analysis may constitute a guideline to further the understanding of how the brain represents and organizes the external world.

5.2 Single-cell and collective behavior

In the previous section we discussed code as the "software" of brain function, the program that determines what the brain has to do. A different meaning of code is related to the "hardware" or physiological implementation of the brain program. In this regard, one of the more debated issues is whether the relevant unit of processing in the brain is the single unit or neural ensembles. We will interpret the results on syllabic representation within the framework

of both theories.

In the first place, is it possible to interpret these results as a case of single-cell coding, i.e. that the syllabic maps reflect the responses of "complex cells", placed in a hierarchy of neuronal relays that are increasingly selective for song-like features of the stimuli? It is usual to describe neurons involved in sensory processing as constituting some sort of a continuum of response complexity, from "simple" cells that respond non-selectively to a variety of stimuli parameters, to "complex" cells that are sharply tuned to a very specific combination of parameters. The actual "optimal" feature combination for a given complex cell appears to depend on the ecology of the species studied. For example, as discussed in [Gross 92] about visual responses in the primate inferior temporal (IT) cortex,

> *"With the exception of face- and hand-selective cells, there is no evidence for IT cells that are selective for visual objects such as fruit, tree branches, monkey genitalia, features in the monkey's laboratory or natural environment or any other object.(...) Perhaps as distinguishing among apples, for example, is so much less important for monkeys than distinguishing among faces, apple-detecting cells may be so rare as to make their discovery highly improbable, but unless evidence for them appears, it seems more parsimonious to assume that they do not exist".*

Applied to NCM, this view would postulate the possible existence of

complex cells selective for the species's own repertoire. But what would happen if NCM's auditory responses were exclusively the result of just a large group of neurons whose activation depends on precise stimulus conditions? In particular, what would be NCM's response to natural and artificial stimuli if the system could be reduced to a collection of complex cells? Such a view would predict that more cells would respond to a natural whistle than to a synthetic one, since the latter is an impoverished version of the former in terms of frequency and amplitude envelope. The same should be true for natural whistles vs. guitar notes, unless one were to argue that the canary's hearing apparatus is more tuned to guitar notes than to canary syllables.

Indeed, the actual responses in NCM are diametrically opposite to the prediction above: responses are bigger and stronger (more cells with broader spatial and labeling distributions) for stimuli with fewer features. Especially compelling is the fact that the average number of NCM cells that respond to a single guitar note comes close to the number responding to a whole canary song (821 ± 77 and 996 ± 92 respectively). The purpose of this exercise is to argue that even if we conceded that one can unambiguously define the "receptive field" properties of a single cell, with some complex kernel that defines its selectivity to stimulus parameters, the collective behavior of large ensembles of cells cannot be trivially extrapolated from this knowledge. In fact, an interpretation based on single cell selectivity for complex combinations of stimulus parameters does not seem to account for most of our results.

In particular: **(a)** it does not explain the habituation experiment: the region activated in the reinduction group is strikingly different from that activated by the same combination stimulus before habituation, instead of revealing some sort of "combination" cells; **(b)** it does not explain the systematic non-linearity (moreover, subadditivity) observed whenever a stimulus is broken into smaller pieces; **(c)** it does not explain the massive topographic overlaps across stimuli. The evidence against a single-cell interpretation is clear and compelling.

5.3 Ensemble Coding

Can the syllabic maps in NCM be understood as a case of ensemble coding? The common understanding of ensemble coding is illustrated in the definition of population vector given in [Georgopoulos 95] regarding neuronal responses in the primate motor cortex:

> *"The broad directional tuning of single cell activity indicates that a given cell participates in movements of various directions; from this result, and from the fact that preferred directions range widely, it follows that a movement in a particular direction will engage a whole population of cells. A unique code for the direction of movement (...) regarded this population as an ensemble of vectors. Each vector represents the contribution of a directionally-tuned cell (...). The weighted vector sum of these*

neuronal contributions is the population vector:

$$P(M) = \sum_{i}^{N} w_i(M)C_i$$

where M is the direction of movement and C_i is the preferred direction of the i-th cell ".

The main assumption made by A. P. Georgopoulos is that responses arise from the linear summation of independent single-cell responses. Ensemble coding, in that sense, means basically that any given stimulus will activate large numbers of neurons, which will add their individual responses in a unique population response. Because of its linearity, the system can be reduced to the sum of its components, i.e., the properties of the ensemble are an expression of equivalent properties of the cells themselves, the difference between them being just a question of number.

However, the results suggest an altogether different notion of ensemble coding. As discussed in the previous section, the most parsimonious explanation appears to be that syllabic representations in NCM are encoded by sets of highly interacting neurons, rather than by cells that function in parallel as independent filters of complexity. Thus, the properties of the responses (such as non-linearity and partial overlaps) revealed in our study are properties of the global patterns of activation, and not of the individual cells that participate in them. In other words, they should be regarded as properties that emerge from the collective behavior of neurons within NCM, qualitatively different from individual neuronal properties. The evidence for such a

notion of ensemble coding in the brain is recent and still very scarce. One of the best examples comes from studies of odour coding in the honeybee brain [Joerges 97]:

> *"Mixtures evoked patterns that were combinations of the single odorant responses. These combinations were not fully additive, however, indicating inhibitory effects on single glomeruli. Such interactions could be crucial for the formation of singular codes for complex odour blends. (...) Consequently, the observed non-linear mixture effects could be the result of interactions at the receptor level, or of network interactions in the antennal lobe, or both."*

The results presented in the previous chapter provide strong and quantitative evidence in favor of such a coding scheme in the brain of a higher vertebrate, in that non-linearity was observed and measured for several stimuli combinations at cellular resolution. As a consequence, each stimulus, within a particular context, recruits a subset of the total pool of neurons in a unique manner. Moreover, the set of responses preserves, to a large extent, the structure of the set of stimuli, a basic requirement for a brain code.

5.4 Toward a model of auditory representation

5.4.1 Programming the syllabic maps?

In Section 1 of this chapter we advanced the hypothesis of a mechanism of self-organization underlying the phenomena observed in the canary brain, and as a general strategy of neural processing. A strong evidence in support of this hypothesis is provided by preliminary results on the ontogenesis and adult plasticity of whistle maps. The basic questions that one wants to answer are: *to what extent is the construction of syllabic maps determined by the auditory input the animals are exposed to, and how much can we shape these maps?* The last question is indeed related to the general problem we posed at the end of Chapter 1: can we program a biological system?

In order to answer these questions we investigated the changes in the representation of canary whistles in juvenile and adult canaries upon manipulation of the auditory input. The results are summarized in Figure 5.5: Panel **a** displays the map elicited by a natural 2.2 kHz canary whistle in the NCM of adult (1 year old) canary females; Panel **b** displays the map elicited by the same stimulus in adult animals housed during 3 months in a zebra finch aviary. Zebra finch is a songbird species whose song shows striking differences with canary song, most notably in the frequency components of the syllables (typically having many more harmonics) and also in the score, being more 'rhythmic' than the 'melodic' canary song (Figures 5.2 and 5.3).

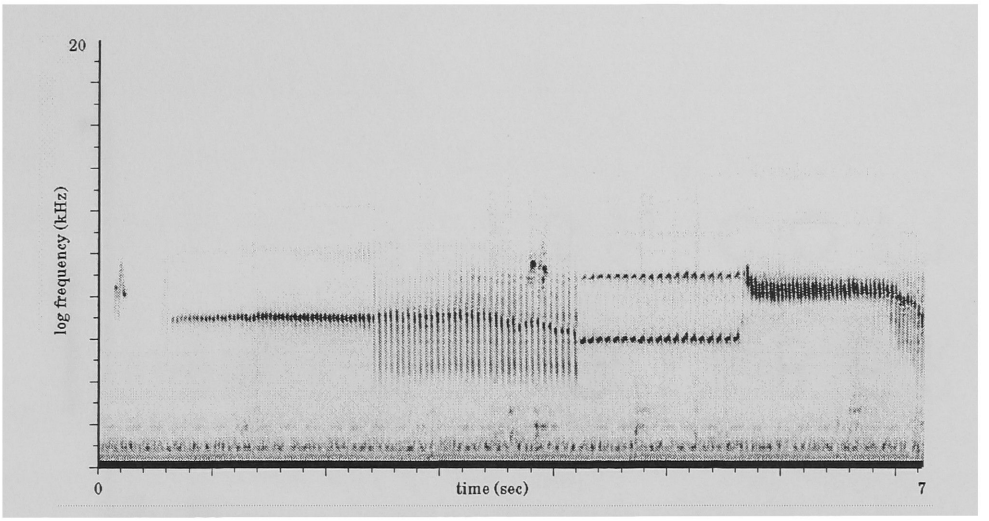

Figure 5.2: Sonogram of a few seconds of a typical canary song

Panel **c** shows the map elicited in the NCM of juvenile canaries (33 days old) normally risen in a canary aviary. Finally, Panel **d** shows the map elicited in juvenile animals which, 33 days after hatching, where housed for 20 days in a zebra finch aviary.

There are clear differences between the normal adult and juvenile maps: first, they differ in the number of cells and the clustering (Figure 5.5, panels **a** and **c**). This difference could be understood as simply an overall increase of the activity in the juveniles, as part of the process of formation of the maps. However, panels **a** and **c** show a striking difference in the topography of the maps: the juvenile, although more extended, seems to elicit little activity in the region that correspond to the adult map 'center of mass'. One would be tempted to conclude that *not even the basic topography of the map* has developed in the juvenile brain, but unfortunately we will have

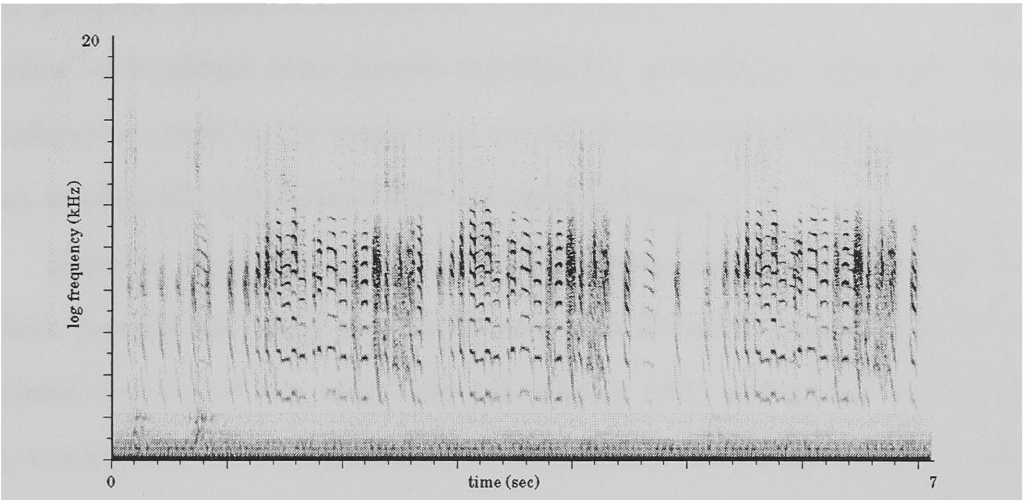

Figure 5.3: Sonogram of a typical zebra finch song

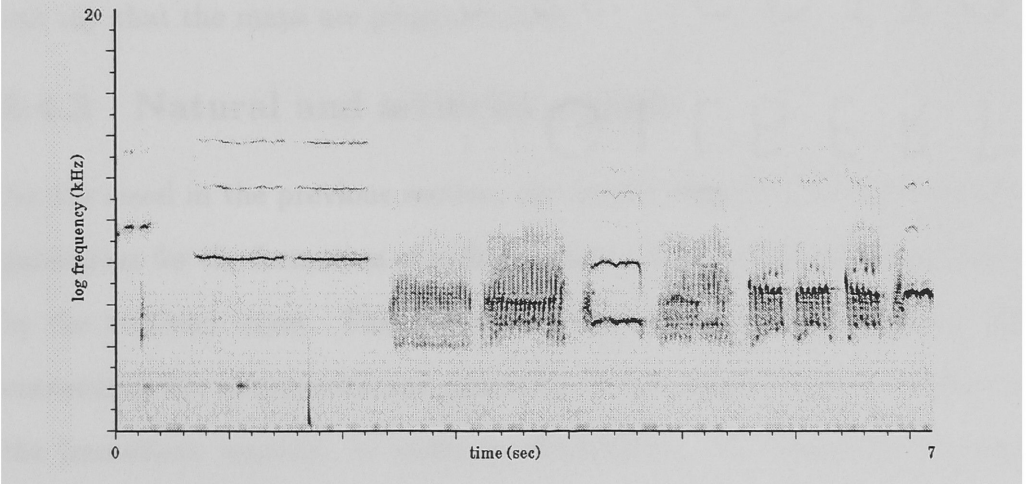

Figure 5.4: Sonogram of the song produced by a juvenile male housed with zebra finches

to postpone definitive conclusions in this regard until further investigations allow us to obtain more precise topographic information. Moreover, these differences could be the result of a normal development of the maps, having no relationship whatsoever with the auditory input.

However, in support of the initial hypothesis, panels **b** and **d** show that both juvenile and adult animals housed with finches change drastically the global activity of the maps (see also Figure 5.6), and in the case of the juvenile, also the topography. The preliminary conclusion is that syllabic maps can be strongly altered by the manipulation of the auditory input, dynamically allocating neural space. We are currently analyzing the song of the male canaries that were housed with finches, which appears highly distorted with respect to normal canary song (Figure 5.4). In a sense, we can say that the maps are *programmable*.

5.4.2 Natural and artificial stimuli

As discussed in the previous section, the results suggest a mechanism of organization for the formation of syllabic maps driven, at least to some extent, by the auditory input. This hypothesis has several consequences not only concerning the ontogonesis and plasticity of the maps, but also in terms of the immediate response to auditory stimulation. To exemplify this issue, we will concentrate ourselves on the phenomenon of spatial clustering of the responses to natural (as opposed to artificial) stimuli described in Section X of the previous chapter.

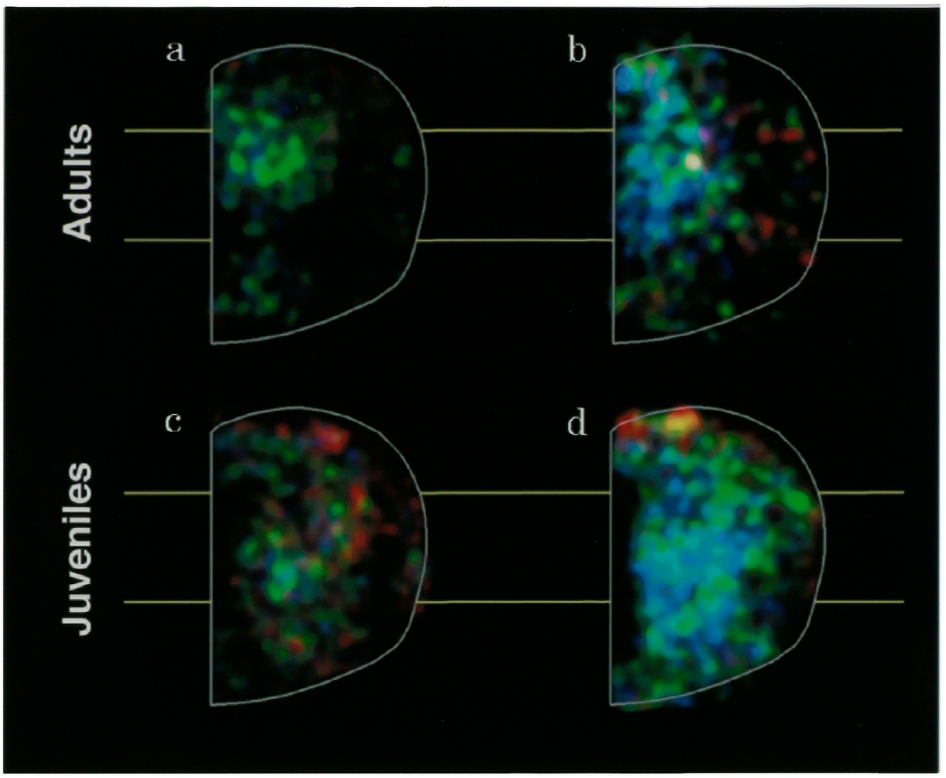

Figure 5.5: Maps elicited by a natural 2.2 kHz canary whistle in the NCM of juvenile and adult canaries. Panel **a** corresponds to normal adult animals, Panel **b** to adult housed with zebra finches for 3 months, Panel **c** to juvenile animals (33 days), and Panel **d** to juvenile animals housed with zebra finches for 20 days

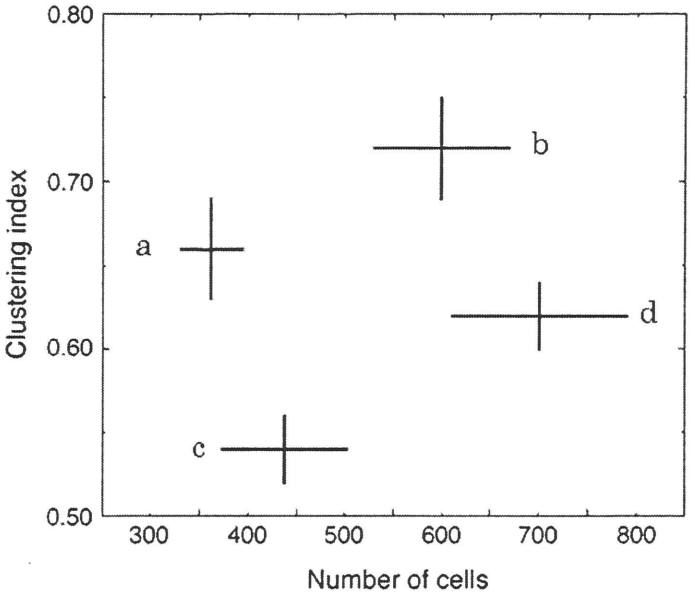

Figure 5.6: Clustering and number of cells corresponding to the four cases shown in Figure 5.5. Bars display STDs

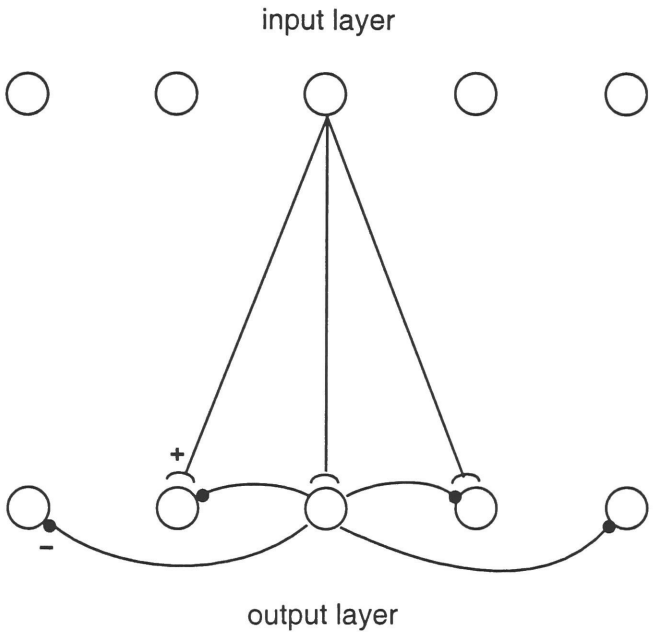

Figure 5.7: Lateral inhibition architecture: the input layer projects excitatory feedforward connections, and each unit in the output layer inhibits its neighbors

We choose the Kohonen map previously introduced as a general algorithm of unsupervised learning, for which the physiological implementation seems plausible. In fact, the lateral inhibition architecture, as part of the general mechanisms of winner-take-all, seems to be pervasive to many brain areas. We will show that an interpretation based on this architecture can shed light on two specific aspects: how natural stimuli can elicit more clustered responses than artificial ones, and what could the consequences of clustered responses be for brain function.

Without loss of generality for the present purpose, we will assume rate-coded linear units. The input layer is represented by the activity of n units,

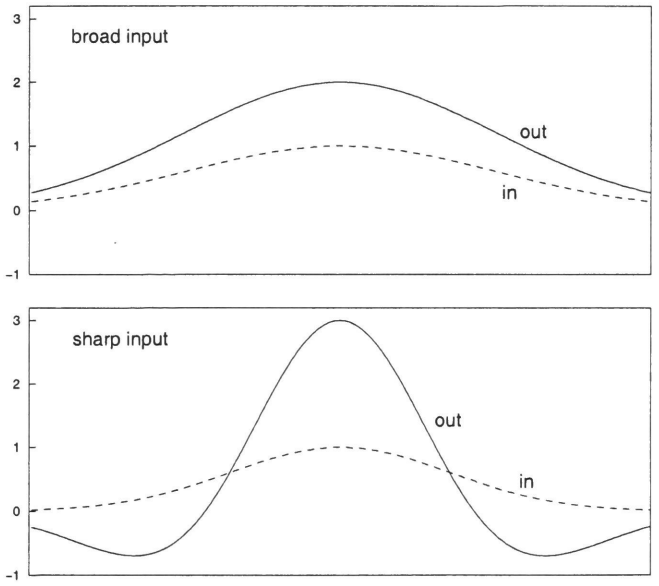

Figure 5.8: Continuous model of lateral inhibition. The upper panel shows the response of the output layer to a 'broad' input; the lower panel corresponds to the response to a 'sharp' input with the same peak value

ϕ_i; at the output layer, each unit ψ_i receives input from each input unit with the corresponding synaptic weight, $S_{ij}\phi_j$, and from the lateral connections, $L_{ij}\psi_j$. This can be summarized as:

$$\vec{\psi} = S\vec{\phi} + L\vec{\psi} \tag{5.1}$$

where S is the $m \times n$ feed-forward synaptic matrix (m is the number of output units) and L is the $m \times m$ lateral synaptic matrix ($L_{ii} = 0$). The algorithm can be seen to operate in cycles, where each cycle implements an instance of Equation 5.1; this process guarantees the convergence to a winner and to zero activity for the remaining units [Kohonen 96]. In spite of its precise physiological implementation (for a discussion, see [Kohonen 96], the goal of the learning process is to maximize the activity of the winner unit. This is equivalent to maximizing the dot product

$$y_w \sim S_{wj}\phi_j = \vec{S}_w \cdot \vec{\phi}$$

which can be achieved by "shifting" $\vec{S}_w$ to make it parallel to $\vec{\phi}$:

$$\Delta\vec{S}_w \sim (\vec{S}_w - \vec{\phi})$$

This means that, after training, the presentation of a particular example in the training set - a 'natural' stimulus - will elicit a *sharp* activity profile in the output layer, exciting the corresponding winner and a few next neighbors. What happens then when an 'artificial' stimulus (i.e., an example drawn form a different statistical ensemble) is presented to the network, after it

has been trained? Typically one can assume that this stimulus will span several components of the set $\{\vec{S}_i\}$, and therefore will look 'broad', exciting more neighbors than the natural ones. We can understand qualitatively the response of the output layer to the new stimulus by rewriting 5.1 in the following way:

$$\psi_n \sim \phi_n - \psi_{n+1} - \psi_{n-1}$$

where n is the spatial index of the units and the synaptic matrix is assumed to be the identity. Therefore,

$$3\psi_n \sim \phi_n + 2\psi_n - \psi_{n+1} - \psi_{n-1}$$

which in the continuous limit yields $\psi(x) \sim \phi(x) - \nabla^2\psi(x)$, and in its turn can be expanded as $\psi(x) \sim \phi(x) - \nabla^2\phi(x) + \nabla^4\phi(x)....$ Therefore, the lateral inhibition architecture determines, to first order, an input-output relationship of the form

$$\psi(x) \sim \phi(x) - \nabla^2\phi(x)$$

from which is evident that a 'sharp' input will elicit a stronger and more restricted response than a 'broad' one, as can be appreciated in Figure 5.8. This result provides a parallel to the clustering observed in the response to natural whistles, and can also explain why a global increase in the activity, if 'flat' (as in the case of artificial whistles), can result in a weaker response in the next processing station.

The rate code assumption is, however, a particularly crude approximation to the problem of auditory coding, where time information is of crucial

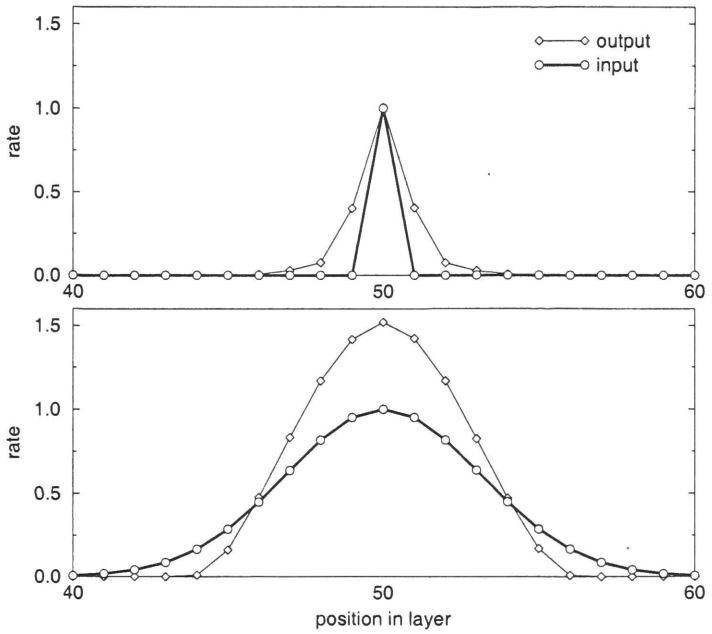

Figure 5.9: Average firing rate across the output layer for 'sharp' (up) and 'broad' (low) inputs in the continuous time model

importance. With the purpose of understanding the temporal behavior of a processing layer with lateral inhibition, we will model it in the following way:

- the units are noisy integrate-and-fire model neurons, essentially identical to those described by Equation 1 of Chapter 3;

- the input layer projects a diffuse connection, so that the synaptic weights form a distribution

$$S_{ij} = Se^{-\alpha(i-j)^2}$$

- the lateral connections form a 'Mexican hat' pattern of synaptic weight,

$$L_{ij} = L[e^{-\beta(i-j)^2} - e^{-\beta(i-j-\gamma)^2} - e^{-\beta(i-j+\gamma)^2}]$$

  that is to say, exciting the immediate neighbors and inhibiting distant ones;

- the lateral connections have a fixed transmission delay τ, i.e. the synapse is excited τ time units after the unit crossed the threshold and spikes;

- pre-synaptic activation induces a fixed increase (excitatory) or decrease (inhibitory) of the post-synaptic potential, equal to the synaptic weight.

The results of simulating this network are shown in Figures 5.9 and 5.10, where a constant current is applied to the input layer, with a sharp or broad

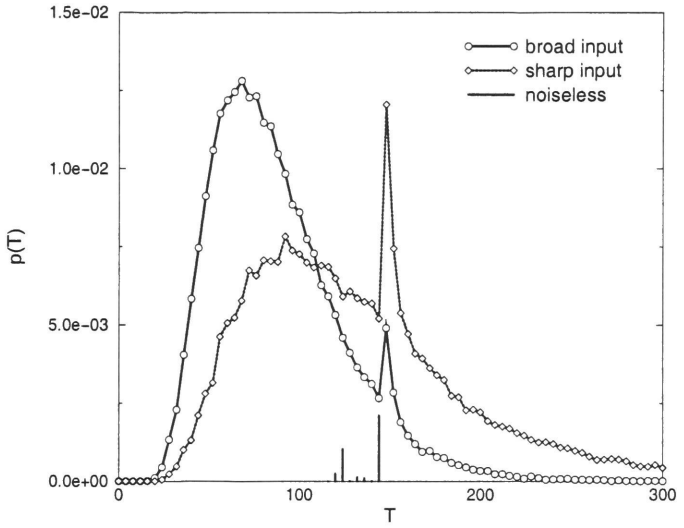

Figure 5.10: Probability distribution of ISI for 'sharp' (diamonds) and 'broad' (circles) inputs. The short trace (lines) corresponds to the noiseless distribution (not normalized)

distribution across the 'receptors'. Both layers consist of 100 units, and like in the rate-coded case, sharp and broad inputs have the same peak value.

Figure 5.9 displays the average firing rate across the output layer. As opposed to the rate-coded simulation, the broad input elicits a stronger response than the sharp one, and therefore the clustering of the sharp input is simply the result of less activity. However, sharp and broad stimulations affect very differently the temporal structure of the response. This can be appreciated in Figure 5.10, where the traces correspond to the distribution of inter-spike intervals (ISI) of the winner unit (n=50 by construction) for the sharp input (dotted line with diamonds), and for the broad input (solid with circles). For comparison, the distribution of ISI for noiseless units is also shown (solid lines, not normalized). Clearly, the sharp input preserves a peak near the most probable interval of the noiseless situation (corresponding to the time to integrate the constant input), which has almost disappeared in the broad input ISI. This can be understood as follows: the structure present in the noiseless case is an indication of the existence of lateral 'rebounds' of activity following a spike in the winner; in the noisy case these rebounds are transformed into a stochastic background, whose total power increases when the broad input activates more neighbors, therefore washing out the initial peak.

We claim that, despite the simplicity of this model, this basic mechanism will strongly affect the processing of 'natural' as opposed to 'artificial' stim-

uli by networks with a Mexican hat architecture. In particular, it is relevant to point at the growing literature on the effect of the temporal structure of calcium excitation on various cellular mechanisms related to neuronal plasticity. We can mention two examples amongst them: in [De Koninck 97] it is shown that the autophosphorilation of CaM Kinase II is modulated by the amplitude and duration of individual calcium spikes; and in [Dolmetsch 98] that calcium oscillations reduce the effective threshold and provide frequency-encoded specificity for the expression of different transcription factors.

5.5 Outlook

A series of questions arise naturally from the previous discussion on coding, plasticity and modeling, which we hope will guide our future research:

(i) How much information is necessary to reliably reconstruct the behavior of the brain? The compromise between the collective aspects of brain function and the complexity of the individual neuron must reach an equilibrium. This issue will become crucial in the near future, as new techniques can provide fantastic amounts of data at both ends of the equation.

(ii) Related to the previous question: how many neuronal internal states have to be considered, and how do they affect short term (the immediate processing of information) and long term (plasticity or neural reorganization) responses? In particular, the existence of these internal states should be considered in network models as an essential element.

(iii) It is clear that the trilogy stimulation-network architecture-genomic activation has to be considered as a unity, determining and depending of each other. The spacio-temporal features of the input (be it sensory or generated by other brain areas), the local architecture of the brain area in study, and the state of genomic activation (and probably, the state of several relevant enzymes) must be considered simultaneously, both in experimental and theoretical terms.

Epilogue

As it may have become apparent, the primary intention of this work has been to generate new questions and to provide a new view into biological phenomena. Although in research good intentions are not enough to produce good work, it has been done with passion and joy. It is the hope of the author to have contributed a small portion to the advancement of the understanding of the brain, the ultimate frontier, and the site, of human knowledge.

Part IV

Addenda

The two Addenda presented here are centered in aspects of molecular computation and the limitations imposed on it by physical constraints, in particular thermal noise and the need to abide by the laws of thermodynamics. In Chapter 1 we will introduce the concept of molecular computation, and in particular DNA based computers. We will show how a graph theoretical problem can be encoded using oligonucleotides, and discuss the potentialities and limitations of this approach. Motivated by the problem of advancing the computation in the context of imperfect molecular interactions, we will present in Chapter 2 a kinetic model of the Polymerase Chain Reaction (PCR), a key component of many DNA based computer schemes. We will show that the sensitivity to the initial conditions leads to non-trivial limiting behavior of the statistics of the output.

Chapter 1

DNA computers

Traditionally, (formal) logic is concerned with the analysis of sentences or of propositions and of proof with attention to the form in abstraction from the matter. This distinction between form and matter is not easy to make precise immediately, but it may be illustrated by examples.

A. Church, Introduction to Mathematical Logic

1.1 DNA Based Molecular Computation

As it we discussed in Part I/Chapter 1, computation with biomolecules is a theoretical possibility. However, until L. Adleman performed a computation with molecules of DNA [Adleman 94], no real implementation had been demonstrated. Following this landmark experiment, there has been theoretical work on variations of the similar computational paradigm [Lipton 95, Reif 95, Beaver 95] but no experimental review of the promises and complications of DNA based computing. In this chapter, we will describe an experiment in which these issues are explored, focusing on heteroduplex formation

during the polymerase chain reaction (PCR) as one critical complication to large scale DNA computing [Kaplan 98].

Why choose DNA for molecular computation? It is clear that many of the challenges of manipulating information on a molecular level have been solved by biological systems; moreover, computationally intractable problems of molecular design have been solved using nucleic acids by the technique of *in-vitro* evolution [Joyce 94, Szostak 93]. By pointing out that today's biotechnology allows us to bring over a billion years of evolutionary development to molecular computation, Adleman has allowed us to foresee a day when molecular computing will be a reality.

Adleman's key innovation was finding a method of encoding an NPC problem in DNA sequences: finding a Hamiltonian Path on a directed graph, similar to the Traveling Salesman problem, and transformable into it (both are NPC problems). A directed graph is a set of vertices connected by edges on which it is only possible to move from one vertex to another if there is an edge connecting them in the correct direction (*e.g.* Fig. 1.1a). A Hamiltonian Path passes through the graph visiting every vertex exactly once; the path must start at vertex S and end at vertex E. The computer is realized by using single stranded 20 nucleotide sequences of DNA to represent each vertex and each edge. The edge sequences are chosen so that they will hold two vertices by base pairing (Fig. 1.1b). By making a mixture of all edges and vertices, we imagine that all possible paths will randomly

assemble; the computational task is to sort through these possible paths for Hamiltonian Paths. Another way to describe this is to call the vertices words and the edges (grammatical) rules which specify which sequence of words are legal. The original idea was to utilize the intrinsic parallel power of DNA to obtain an algorithm able to solve the Hamiltonian Path in polynomial time, and therefore demonstrating not only that molecular computation is possible, but that it can provide unprecedented computational power. We will see that physical constraints and technological limitations conspire to curtail the capabilities of DNA computers.

1.2 PCR as an amplifier

Our algorithm is to generate a complete language and then sort through it for sentences which represent Hamiltonian Paths. This scheme is generally applicable in the sense that many problems can be solved by generating all possible solutions and then sorting through them for the correct answer. In the course of repeating the previous experiment [Kaplan 96], we found that the major difficulties were at the beginning, in assembling a complete language of paths and in copying a subset of that language by PCR. In this paper we examine apparently oddly sized DNA produced when copying the language and describe one way in which this DNA is produced by PCR.

The polymerase chain reaction is an important tool for DNA based computation. In fact, PCR, as a DNA amplifier, will almost certainly be a key

component of any large scale DNA computer. In PCR, the sample is heated to separate or denature the double stranded template and then each of the two strands is copied. Each time this thermal cycle is repeated, both strands and the copies will be duplicated - therefore the number of copied grows exponentially with the number of thermal cycles. The copying is initiated by the hybridization, of one of two "primer" oligonucleotides (Fig. 1.1c). The assumption that by manipulating a surplus of primer to template we can neglect template/template interactions turns out to be incorrect when amplifying heterogeneous populations of DNA.

As part of our efforts to reproduce Adleman's experiment, we did some experiments on a simple graph (Fig. 1.1a). In the simple graph, the Hamiltonian PAth ($S12345E$) is the only path. In the next two graphs, SG+13 and SG+35, we add a single edge making possible a single non-Hamiltonian Path (for SG+13 it is the path $S1345E$, and for SG+35 it is $S1235E$). To begin the calculation, in a 30 μl reaction, we mix together 50 picomoles of single stranded oligonucleotides representing each edge and each vertex of the graph, and then we allow the oligonucleotides to hybridize in the presence of the enzyme T DNA ligase which forms permanent paths by linking the sugar phosphate backbone of the oligonucleotides (Fig. 1.2 lanes 4-6) [Adleman 94] [1]. After a two hour incubation at 37 °C, 5 μl is transferred

[1] A 94 °C mixture of edges and vertices containing 50 pm of each oligonucleotide was cooled to 37 °C and Promega DNA ligase buffer containing 10 units of T4 DNA ligase was added.

to a 50 μl PCR reaction and paths stretching from S to E are copied using oligonucleotides S and the complement to E as primers [2] (see Fig. 1.1c). The PCR product is examined by polyacrylamide gel electrophoresis [3] (Fig. 1.2 lanes 1-3). On observing the ligation product on this gel (lanes 4-6), we find a continuous distribution of electrophoretic mobilities, not a set of clear bands corresponding to the discrete numbers of oligonucleotides. The continuous distribution may be due to the formation of complex structures, which may involve folded DNA, complexes between strands, or incorrect ligation products. On examining the PCR product (lanes 1-3) we note that the simple graph (lane 3) produces a single strong band at 140 bp (7 steps times 20 bases/step) as expected. In the other graphs (SG+13 lane 5, SG+35 lane 1), in addition to the expected bands at 140 bp (7 steps) and 120 bp (6 steps), we find a striking band at about 190 bp. This band is surprising because the only way to generate a path longer than 7 steps is to have an

[2] 50 μl PCR reactions were performed in a buffer of 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001 % gelatin, 100 μm each of dATP, dTTP, dCTP, and dGTP, 50 pm of both primers, and 2.5 units of recombinant Thermus aquaticus DNA polymerase. The complex temperature cycle was: 94 °C for 5 min.; 35 cycles of (94 °C, 30 s; 55 °C, 30 s; 72 °C, 60 s); 72 °C for 5 min. The reaction was then cooled to 4 °C in a Perkin Elmer model 9600 thermal cycler.

[3] Native polyacrylamide gels were made by casting a mixture of 6 ml of 40 % 19:1 Acrylamide:bis-acrylamide, 30 ml H_2O, 4 ml 10X TBE (1 M Tris, 0.9 M Boric acid, 10 mM EDTA), 270 μl of 10 % Ammonium persulfate and 50 μl of TEMED (N,N,N',N'-Tetramethylethynediamine) between glass plates separated by 1.5 mm. A teflon comb was inserted between the plates to make loading wells. Gels were allowed to polymerize for at least 2 hours. Samples mixed with a 6X loading buffer of 10 % Ficoll in 50 mM Tris (pH 8.3). After loading gels, equilibrated for at least 30 minutes prior to electrophoresis for 50 min. in a running buffer of 1X TBE at 13 V/cm and 25 °C. Gels were visualized using a UV transilluminator after a 10 min. soak in 1 μM of ethidium bromide.

edge pointing backwards; no such edge is included in Fig. 1.1a. Further, 190 bp corresponds to 9.5 steps. We do not know how to interpret a half step. We decided to explore this 190 bp band as a prototype of the unexpected bands seen in experiments on DNA computation.

We isolated the DNA in the anomalous band by cutting it from the gel, soaking the gel slice in water overnight, and amplifying DNA in the soak water in PCR. Although we expected to see only a band at 190 bp, the PCR product included all three original bands (120, 140, and 190 bp). By varying the amount of DNA used as input to PCR, we found that at low concentrations, only the short bands are produced, but at higher concentrations, all three bands appear (Fig. 1.4 lanes 1-4). To rule out the possibility of mis-priming during PCR, we performed a complete set of optimization reactions, varying the pH from 8.5 to 10, the concentration of Mg^{++} from 1.5 to 3.5 mM with both 0 and 10 % DMSO [Invitrogen]. All three bands (120, 140, and 190 bp) always appeared together. To explore the possibility that an illegal or ungrammatical path was responsible for the 190 bp band, we performed PCR using many combinations of vertices and their complements as primers. No coherent patterns emerged.

We then examined the PCR product on a denaturing gel [4] on which DNA separates into single stranded molecules (Fig. 1.3). On the denaturing gel,

[4]Denaturing gels were run as native gels except that the gel contained 7 M Urea, 12 % acrylamide, the loading buffer was a 50/50 mix of sample and formamide, the electrophoresis apparatus was heated to 55 °C, the electric field was 21 V/cm and the gel was stained with a 20 min. soak in 1μM SYBR Green 1 (FMC).

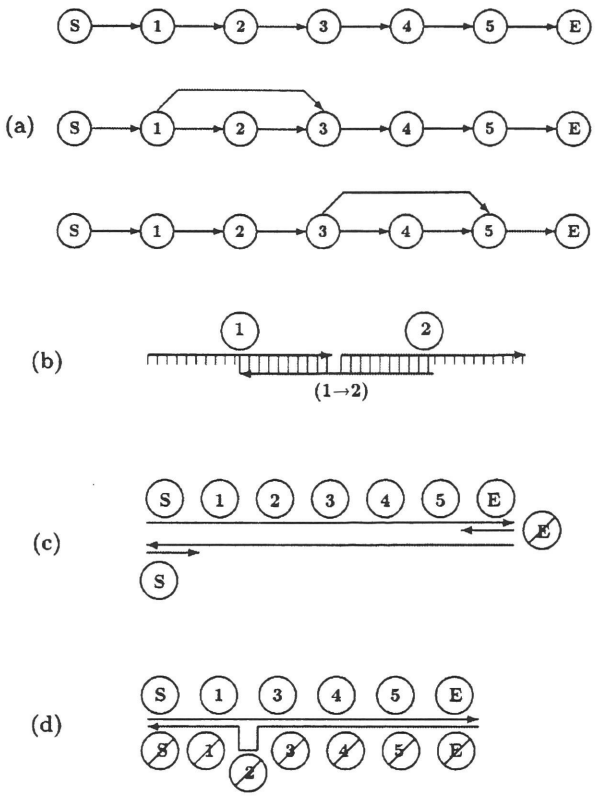

Figure 1.1: (a) By solving the Hamiltonian Path on these three graphs, the Simple Graph (SG),SG+13, and SG+35, we have been able to identify one source of surprising bands in a DNA computation. (b) The vertices of the graph are represented by single strands of DNA. Vertices are held together by annealing to edges whose sequence matches the end of one vertex and the beginning of the next. This structure is the template for ligation. (c) To amplify sequences beginning with S and ending with E, we perform PCR using S and the complement to E as primers (in the figure the complement is denoted by a slash through the symbol). (d) The heteroduplex structure formed by 120 and 140 nucleotide strands annealing during PCR. This structure migrates slower than 120 and 140 bp double stranded DNA.

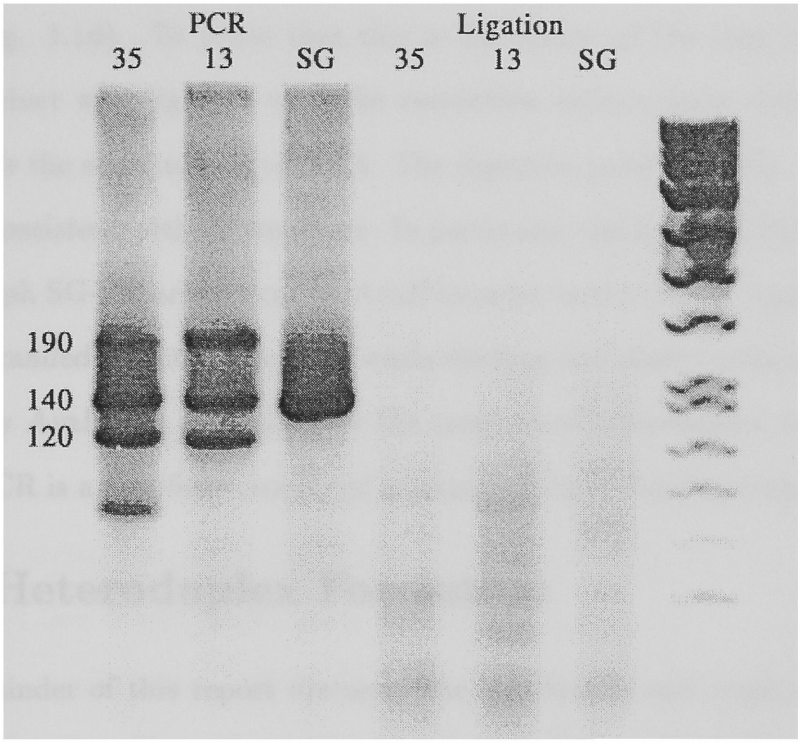

Figure 1.2: Ligation and PCR of the three graphs in Figure 1.1a. While the PCR product of SG (lane 3) displays only the expected band at 140 bp, the products of SG+13 (lane 1) and SG+13 (lane 2) display bands at 120 and 140 bp as well as a surprising band at 190 bp. Note that the ligation products (lanes 4-6) do not migrate in discrete bands.

we found only 2 bands (120 and 140 bp). From this result, we inferred that the DNA in the long band was a hybrid of complimentary 120 and 140 bp DNA (Fig. 1.1d). To prove that this is the source of the long band, the PCR product was digested with the restriction endonuclease *AvaII* which cuts inside the sequence for vertex 4. The digestion products (Fig. 1.4 lanes 5-7) are consistent without structure. In particular, the long and short bands in the graph SG+35 are not cut by *AvaII* because vertex 4 never appears as a *double stranded* object in this DNA while the long and short bands in SG+13 *are* cut by *AvaII*. We conclude that the creation of heteroduplex structures during PCR is a significant source of unexpected bands during computation.

1.3 Heteroduplex Formation

The remainder of this report discusses the significance and implications of heteroduplex formation in computation and in biology. Biologists use heteroduplex formation to identify insertion or deletion mutations [Triggs-Raine 90, Lilley 95]. PCR amplification of the gene of interest is performed in the presence of a non-mutant allele and the results are analyzed by gel electrophoresis. As the normal and mutant allele can defer by one or two bases, they are difficult to distinguish directly on a gel. The heteroduplex's single stranded bulge results in a vastly different mobility creating a distinct, distant band which is a clear signature of mutations [Lilley 95].

In the context of computation, the phenomenon of heteroduplex forma-

tion requires careful consideration. First, heteroduplex formation will be ubiquitous in DNA based computation in which PCR is used as an amplifier. In this simple graph there is a simple heteroduplex; in a more complicated graph (*e.g.* [Adleman 94]) there is an enormous set of possible heteroduplex structures. Next, although it is the best available tool, gel electrophoresis is not a good tool for analyzing DNA. As mobility varies greatly with DNA conformation we should use denaturing or single stranded gels for analysis; we hope that a new technology will replace electrophoresis. Next, there are implications for designing calculations. Heteroduplex formation can occur during the annealing step of each cycle of PCR. If the 3' end of a valid computational string anneals to the middle of another string, DNA polymerase will extend this 3' end producing a new, illegal string while simultaneously destroying a valid string. Extreme caution is required to design protocols so that 3' ends will not hybridize incorrectly and so that broken strings are removed before they contribute their 3' ends to a new, illegal, hybrid string. Finally, it is not necessary to use high concentrations of template or to produce large concentrations of product during PCR. By keeping all concentrations low, we believe that heteroduplex formation and other template/template interactions will be minimized. A good computational design should take advantage of this process to create new algorithms which use relatively infrequent template/template hybridization to *advance* the calculation.

1.4 Discussion and Outlook

It is very clear that several difficulties maim the prospects of Adleman's model of molecular computation, besides those discussed in the previous sections (see also [Ouyang 97]. Perhaps the most serious one is the tradeoff between computational time complexity and the initial amount of DNA required. This can be understood as follows: the intrinsicly parallel nature of the computation implies that an unlimited number of molecules (in principle) participate in the search of the correct solution along the tree of possible solutions, rendering the time complexity simply linear in the size of the problem. On the other hand, these molecules are not akin to parallel processors, as they do not communicate with each other; therefore, the exponential time of NP problems is transformed into an exponential increase in the number of molecules needed to fill the solution tree such that the branch containing the correct solution is explored with non-negligible probability. Nevertheless, the value of this approach has been to show that molecular computation is possible, even when the experimental techniques were not originally designed for this purpose. The pursuit of solutions to the difficulties in finding an effective model of DNA based molecular computation should tighten the bonds of theoretical approaches and traditional molecular biology, in particular to devise ways of using the enzymatic machinery of the cell for computational purposes.

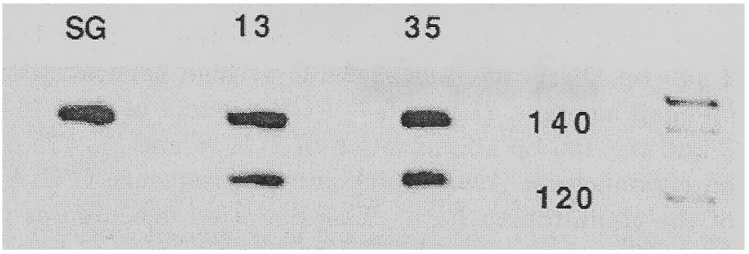

Figure 1.3: Denaturing gel of PCR product displayed in Figure 1.4 (lanes 5-7). When the strands are separated, we see only two sizes of DNA, 140 and 120 nucleotides. This suggests that the DNA in the apparently 190 bp band is actually the heteroduplex structure sketched in Figure 1.1d.

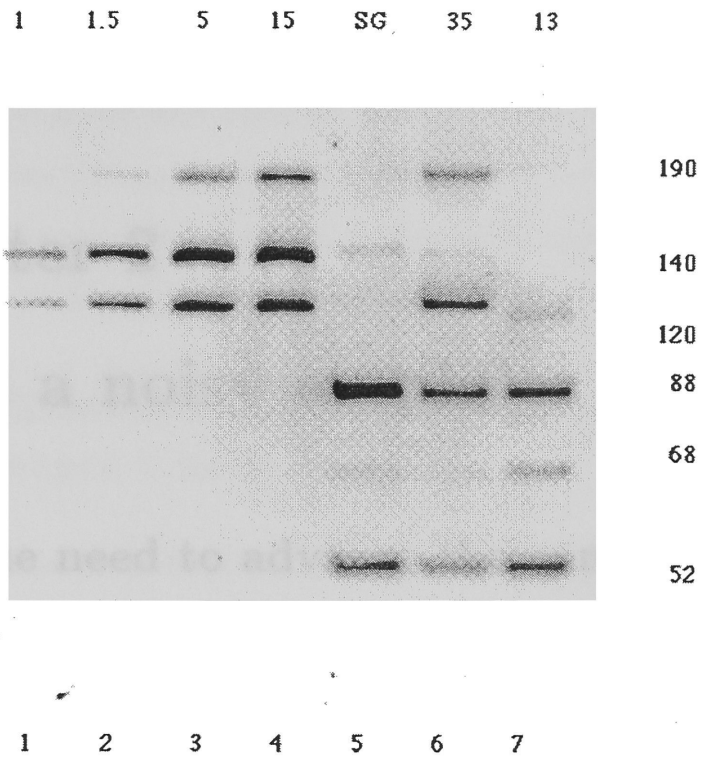

Figure 1.4: Details of the anomalous 190 bp band. (Lanes 1-4) We amplify the 190 bp band with different amounts of 190 bp input (0.5, 1.5, 5 and 15 νl of the water used to soak the gel slice from 190 bp in Figure 1.2 lane 6, were added to a 100 μl 15 cycle PCR reaction). For small quantities of input DNA, the shorter 120 and 140 bp DNA is produced while at high concentrations, the surprising 190 bp band emerges. (Lanes 5-7) PCR product of the 140 bp amplification of SG and the 190 bp amplification of SG+35 and SG+13 digested by the restriction endonuclease *AvaII* which cuts the sequence G'G(A/T)CC at position 88 of the Hamiltonian Path. This digestion decomposes the 140 bp band into pieces 52 and 88 long, the 120 bp band of SG+13 into 52 and 68, the 120 bp band of SG+35 is not cut. The critical result of this digestion is that the 190 bp band of SG+13 is cut while the 190 bp band of SG+35 is not cut because vertex 4 is part of the single stranded loop in the heteroduplex and *AvaII* will not cut single stranded DNA. *Methods:* 20 μl of the PCR product displayed in lanes 4-6 of Figure 1.2 are mixed with 24 units of *AvaII* (New England Biolabs) in the manufacturer's recommended buffer with 0.1 % Bovine Serum Albumin. The reaction is incubated at 37 °C for 16 hours.

Chapter 2

PCR: a noisy amplifier

2.1 The need to advance the computation

Even though Turing's model of computation enforces a "mechanistic" interpretation of computable functions, in actual computational devices, be they electronic or molecular, the different sources of noise (most conspicuously thermal) make it necessary to amplify the amplitude of pulses, the number of oligonucleotide molecules or the concentration of signaling molecules in enzymatic cascades. In the latter case, it has been argued [Magnasco 97] that the existence of logical gates implemented by enzymatic reactions is not sufficient to achieve universal computing power: *global stability* is necessary to advance the computation. In other words, the successive layers of computation should stack or connect with each other reliably. As discussed in the previous section, experimental implementations of DNA computers rely on the amplification of strings coding approximate solutions at different stages of the computation. This amplification process, in its turn, introduces errors

determined by thermodynamic constraints, which given its exponential nature, can have drastic and unpredictable consequences. In the remaining of the chapter, we will explore some of them, bearing in mind the implications for molecular computing as well as technological applications and cellular mechanisms [Stolovitzky 96].

2.2 Efficiency of DNA replication in the Polymerase Chain Reaction

The Polymerase Chain Reaction (PCR) is one of the most widely used techniques in modern molecular biology. It was devised [Mullis 87] as a method for amplifying specific DNA sequences (targets), and the scope of its applications stretches from medicine [Gibbs 90], through *in vitro* evolution [Gestland 93], to DNA based models of computation described in the previous chapter. In spite of its ubiquity in biology, theoretical discussions of PCR are rare. Although kinetic models of the enzyme-mediated polymerization of single-stranded DNA have been reported [Capson 92, Patel 91, Benkovic 87, 91], none of them were applied to model PCR, and only recently a treatment of the rate of mutations arising in PCR has been considered [Sun 95, Weiss 95].

A detailed quantitative kinetic model for the Polymerase Chain Reaction (PCR) is presented here, which allows the prediction of the probability of replication of a DNA molecule in terms of the physical parameters involved

in the problem. The important issue of the determination of the number of PCR cycles during which this probability can be considered to be a constant is solved within the framework of the model. New phenomena of multi-modality and scaling behavior in the distribution of the number of molecules after a given number of PCR cycles are presented. The relevance of the model for quantitative PCR is discussed, and a novel quantitative PCR technique is proposed.

The main object of our study is the probability that one molecule will be replicated in one PCR cycle, the so-called efficiency p. In Section II we present a detailed kinetic model of the polymerization and find p as a function of the physical parameters of the problem. This allows us to discuss the range of validity of the assumption of constant probability of replication, on which statistical considerations have been based [Sun 95, Weiss 95]. Within that range, we apply the theory of branching processes in Section III, to show the existence of new phenomena: the probability density function (*pdf*) of the number of molecules after a given number of cycles of PCR displays scaling behavior, and under some conditions, this *pdf* is multi-modal. In Section IV a novel method for quantitative PCR is presented, based on the statistical considerations of the previous sections. In Section V we summarize our work.

One cycle of PCR consists of three steps. (For a more detailed account of the PCR technique see, e.g., [Sambrook 89].) In the *denaturing* step, the two strands of the parent DNA molecule in solution are separated into single-

stranded (ss) templates by rising the temperature to about $95°C$ to disrupt the hydrogen bonds. In the *annealing* step, the solution is cooled down to approximately $50°C$ to allow the *primers*, present in a high concentration, to hybridize with the ss DNA. The primers are two (different) 20 to 30 nucleotides long molecules which are Watson-Crick complementary to the 3' flanking extreme of the templates. Once the primer-template heteroduplex is formed, it acts as the *initiation complex* for the *DNA-polymerase*[1] to recognize and bind to. This step is crucial for the specificity of the amplification: only those molecules that have sequences complementary to the primers will be amplified. The last step is a polymerization reaction, in which the solution is heated to $72°C$, the optimal working temperature for *Thermus aquaticus* DNA polymerase. This enzyme catalyzes the binding of complementary nucleotides to the template, in the direction that goes from the primer to the other extreme[2]. Notice that if this polymerization proceeded to its end, at the end of the third step we would have twice as many DNA molecules as we had at the beginning of Step 1. These three steps constitute one cycle of the PCR, which is usually 30 seconds to 2 minutes long. The cycles are repeated a number of times (typically 30) by varying the temperature in the solution, in such a way that the DNA molecules that were synthesized in a

[1] Polymerases are interesting pieces of machinery [Yin 95]. They are responsible for the duplication of genetic information (DNA-polymerases) and its transcription into RNA (RNA-polymerases).

[2] The DNA is a *polar* molecule, and the polymerase can only attach new nucleotides to the 3' end of the molecule that is being extended.

given cycle are used as templates in the following one. In this way one gets an extremely efficient amplification mechanism for DNA.

2.3 Kinetic Model

We will represent the last two steps of a typical cycle of PCR by means of a kinetic model. Our species will be the primers (pr, of length L_p nucleotides), the ss DNA (ss, consisting of $L_p + N$ nucleotides), the heteroduplexes (h_i, formed by one complete ss and the partially assembled complementary strand consisting of the primer and the next i nucleotides), the nucleotides (n, which will be considered identical), the polymerase (q) and the heteroduplexes h_i with the polymerase attached to them (qh_i). Denoting by κ_{2j-1} and κ_{2j} the forward and backward chemical reaction rates, the chemical equations are

$$\begin{array}{lll} \text{Step 2} \{ & ss + pr \overset{\kappa_{1,2}}{\rightleftharpoons} & h_0 \\ \text{Step 3} \left\{ \left\{ \begin{array}{rcl} h_i + q & \overset{\kappa_{3,4}}{\rightleftharpoons} & qh_i \\ qh_i + n & \overset{\kappa_{5,6}}{\rightleftharpoons} & qh_{i+1} \end{array} \right\} \right. & 0 \le i \le N-1 \\ \left. \begin{array}{rcl} h_N + q & \overset{\kappa_{7,8}}{\rightleftharpoons} & qh_N. \end{array} \right. & \end{array} \tag{2.1}$$

Other recognizable species might be present in the chain reaction. This can occur because of substitutions, additions or deletions of nucleotides by the polymerase, or because of the presence of sequence-dependent structures. These will not be taken into account in our model, for the sake of simplicity. Assuming that the effects of inhomogeneities in density and temperature are

irrelevant, it is well known that Eqs. (2.1) lead to a corresponding system of first order nonlinear differential equations for the concentrations of the different species as functions of time, which we are not going to write here (see, for example, [van Kampen 81]). In the above reactions, one should assign a given duration to Step 2 and another to Step 3. For the sake of simplicity, however, we shall consider both Step 2 and Step 3 as running simultaneously in the simulations to be presented below. This is a mild simplification which does not alter the conclusions to be drawn.

The definition of the efficiency p implies that it is simply the ratio between the number of ss molecules that were completely replicated at the end of a given cycle and the initial number of ss molecules in that cycle:

$$p(t) = \frac{[h_N](t) + [qh_N](t)}{[ss](0)} \tag{2.2}$$

Figure 2.1 shows plots of the probability of replication p as a function of time t (which is to be interpreted as the duration of Step 3 in a typical PCR cycle), for different polymerization lengths N.

Since to the best of our knowledge the chemical reaction constants κ's have not been measured for Taq polymerase, we have assumed some values for these constants to exemplify the principal characteristics of our model. It should be stressed, however, that the equivalent to some of these constants have been measured for other polymerases such as T4 polymerase [Capson 92], T7 polymerase [Patel 91] and for DNA polymerase I (Klenow fragment) [Benkovic 87, 91]. The values of the chemical reaction constants

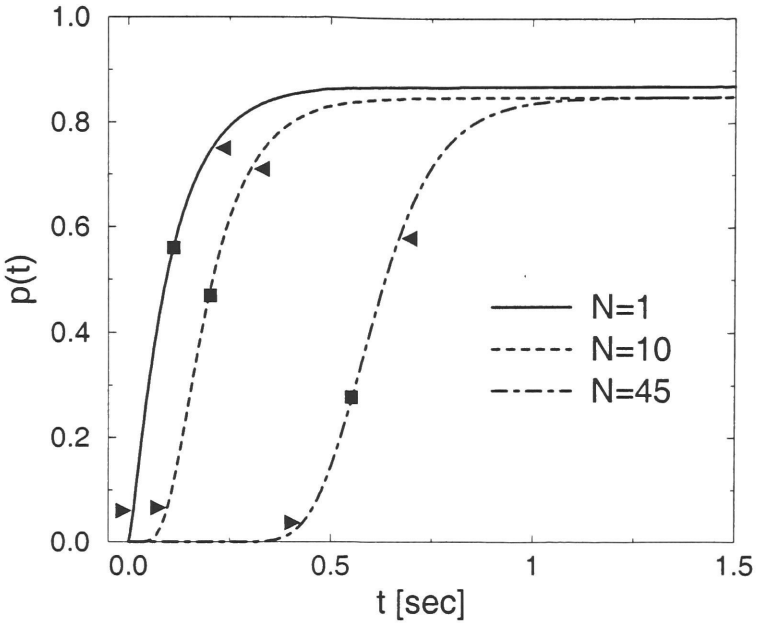

Figure 2.1: Probability of replication $p(t)$ as a function of time t, for different template lengths N (in number of nucleotides without including the primers), arising from a numerical simulation of Eqs. (2.1), with parameters: $\kappa_1 = 10^9 M^{-1}s^{-1}$, $\kappa_2 = 10^{-2}s^{-1}$, $\kappa_3 = 10^7 M^{-1}s^{-1}$, $\kappa_4 = 10^{-3}s^{-1}$, $\kappa_5 = 10^7 M^{-1}s^{-1}$, $\kappa_6 = 15s^{-1}$, $\kappa_7 = 10^9 M^{-1}s^{-1}$, $\kappa_8 = 10^{-1}s^{-1}$, $[pr](0) = 10^{-6}M$, $[n](0) = 10^{-5}M$, $[q](0) = 10^{-6}M$, $[ss](0) = 10^{-11}M$. The square and arrow heads indicate, respectively, the mean rise-time and rise-time width as predicted by the simplified model of Eqs. (2.3)

κ and the initial conditions used in the simulation of Eq. (2.1) are detailed in the caption to Fig. 2.1. The main features of the curves in Fig. 2.1 can be quantitatively understood. It can be observed that the larger N, the flatter the behavior at small times. Indeed it can be shown from the dynamic equations that $p \sim t^{(2N+1)/3}$ for t sufficiently small. The time at which p has reached about half its asymptotic value, as well as the width of the rise-time can be estimated from a further simplification of our model. Assuming that the time constants associated with the backwards reactions in Eq. (2.1) are large enough, and that the concentration of primers, polymerase and nucleotides are sufficiently large that their relative concentrations can be considered as constants (or more precisely as slowly varying parameters) during the process, we can rewrite the reaction as

$$ss \overset{pr}{\underset{\kappa_1}{\rightarrow}} h_0 \overset{q}{\underset{\kappa_3}{\rightarrow}} qh_0 \overset{n}{\underset{\kappa_5}{\rightarrow}} qh_1 \ldots \overset{n}{\underset{\kappa_5}{\rightarrow}} qh_N. \tag{2.3}$$

The time τ needed for this reaction to be completed is simply the sum of the times corresponding to each link of the chain, $\tau = \tau_{\kappa_1} + \tau_{\kappa_3} + \tau_{\kappa_5,1} + \ldots + \tau_{\kappa_5,N}$, where τ_{κ_1} and τ_{κ_3} are the times associated with the first two reactions in Eq. (2.3), and $\tau_{\kappa_5,j}$ is the time associated with the reaction $qh_{j-1} \overset{n}{\underset{\kappa_5}{\rightarrow}} qh_j$. These τ's are independent, exponentially distributed random variables, whose mean values are $\langle \tau_{\kappa_1} \rangle = (\kappa_1[pr])^{-1}$, $\langle \tau_{\kappa_3} \rangle = (\kappa_3[q])^{-1}$, and $\langle \tau_{\kappa_5,j} \rangle = (\kappa_5[n])^{-1}$. Therefore, it can be readily seen that the mean and the

standard deviation of τ are

$$\langle \tau \rangle = \frac{1}{\kappa_1[pr]} + \frac{1}{\kappa_3[q]} + N\frac{1}{\kappa_5[n]}, \tag{2.4}$$

$$\sigma_\tau = \left(\frac{1}{\kappa_1^2[pr]^2} + \frac{1}{\kappa_3^2[q]^2} + N\frac{1}{\kappa_5^2[n]^2} \right)^{1/2}, \tag{2.5}$$

where we used that the variance of a exponentially distributed random variable is the square of its mean. $\langle \tau \rangle$ and σ_τ can be used as estimates of mean rise-time and the rise-time width about the mean for the complete reaction. These estimates are shown in Fig. 2.1. The abscissa of the solid square on each curve corresponds to the value predicted by Eq. (2.4), and the arrow heads indicate the values of $\langle \tau \rangle \pm \sigma_\tau$. It can be safely concluded that Eqs. (2.4) and (2.5), computed from the simplified chain reactions of Eq. (2.3), are good estimates of the mean rise-time and the rise-time width corresponding to the full set of reactions.

The last important feature to be extracted from Fig. 2.1 is the tendency of $p(t)$ towards an asymptote p_∞, which corresponds to the equilibrium of the chemical system. This value is of importance in PCR, and thus it is worth computing it in terms of the parameters of our model. The detailed balance equilibrium conditions for the reactions of Eq. (2.1) demand that $[ss]_{eq}/[h_0]_{eq} = \kappa_2/(\kappa_1[pr]_{eq}) \equiv \alpha_1$, $[h_i]_{eq}/[qh_i]_{eq} = \kappa_4/(\kappa_3[q]_{eq}) \equiv \alpha_3$ (for $0 \leq i \leq N-1$), $[qh_i]_{eq}/([qh_{i+1}]_{eq}) = \kappa_6/(\kappa_5[n]_{eq}) \equiv \alpha_5$ (for $0 \leq i \leq N-1$) and $[h_N]_{eq}/[qh_N]_{eq} = \kappa_8/(\kappa_7[q]_{eq}) \equiv \alpha_7$. On using Eq. (2.2) and the conservation

relation $[ss](t) + \sum_{i=0}^{N}[h_i](t) + [qh_i](t) = [ss](0)$, one obtains that

$$p_\infty = \frac{1+\alpha_7}{1+\alpha_7+\alpha_1\alpha_3\alpha_5^N + \frac{\alpha_3+1}{\alpha_5-1}(\alpha_5^{N+1}-\alpha_5)}. \tag{2.6}$$

For the purpose of computing p_∞ one should know the values of $[pr]_{eq}$, $[n]_{eq}$ and $[q]_{eq}$. As an approximation to these values one can use the initial values of these species at the beginning of the cycle. This approximation will be excellent if this initial concentrations are sufficiently large. The values of p_∞ (computed under this approximation) corresponding to the conditions of the simulations of Fig. 2.1 are 0.87 for $N = 1$ and 0.85 for $N = 10$ and $N = 45$, in perfect agreement with the complete simulation. It is interesting to notice that from direct measurements of $p(t)$ a wealth of information on the rate constants involved in the polymerization reactions can be inferred using Eqs. (2.4)-(2.6).

Of utmost importance in applications of PCR is the number of cycles of PCR during which the amplifying process is exponential. As will be discussed later on, the mean number of molecules $\langle N_{k+1}\rangle$ at cycle $k+1$ is related to the mean number of molecules $\langle N_k\rangle$ at cycle k by the relation $\langle N_{k+1}\rangle = (1+p_k)\langle N_k\rangle$, where p_k is the efficiency during the *k-th* cycle. Therefore the rate of growth will be exponential only when p_k is independent of k. During how many cycles can the system maintain p_k constant? The answer can be found if we think that during these cycles, both the concentration of primers and nucleotides will also be decreasing exponentially, and therefore their concentration at cycle k will be $[pr]_k = [pr]_0 - (1+p)^k[ss]_0$ and $[n]_k =$

$[n]_0 - (1+p)^k N[ss]_0$. The mean rise-time and rise-time width for p at cycle k, $\langle\tau\rangle_k$ and $\sigma_{\tau,k}$, will be given by Eqs. (2.4) and (2.5), with $[pr]$ and $[n]$ replaced by $[pr]_k$ and $[n]_k$ respectively. If the time for the reaction is t, then the maximum number of cycles ν during which p_k can be considered constant will be given, to a first approximation, by the ν that verifies that $\langle\tau\rangle_\nu + \sigma_{\tau,\nu} = t$. This imposes an equation for ν that can be solved numerically. An approximation to this solution is

$$\nu = \min \quad \left\{ \log_{1+p_\infty}\left(\frac{[n]_0}{[ss]_0 N} - \frac{1}{[ss]_0\kappa_5(t-\kappa_3^{-1}[q]^{-1})}\right); \right. \\ \left. \log_{1+p_\infty}\left(\frac{[pr]_0}{[ss]_0} - \frac{1}{[ss]_0\kappa_1(t-\kappa_3^{-1}[q]^{-1})}\right)\right\} \tag{2.7}$$

where log_b indicates logarithm to the base b. Notice that as t becomes larger, the value of ν predicted in Eq. (2.7) tends to a constant independent of t, given by the number of cycles that it takes to deplete the solution of nucleotides or primers, whichever is exhausted first. Although it might be unrealistic for the conditions used in molecular biology, it is interesting to notice that if the nucleotides are the first species to be exhausted, then most of the heteroduplexes will cease to polymerize before reaching the end, with the outcome that there will hardly be any complete double helices formed: in this case the net amplification factor will be close to zero. Figure 2.2 shows the efficiency p_k at cycle k as a function of the number of cycles, for different times of polymerization and $N = 45$ (the other parameters are as in Fig. 2.1), obtained by the integration of the reactions of Eq. (2.1). In this simulation we

concatenated cycles assuming a perfect melting step, which was done by hand by setting $[ss]_{k+1}(0)$ in the cycle $k+1$ equal to $[ss]_k(0) + [h_N]_k(t) + [qh_N]_k(t)$ of the previous cycle and $[h_i]_{k+1}(0) = [qh_i]_{k+1}(0) = 0$, $(0 \leq i \leq N)$. The dynamics of pr and n, on the other hand, was followed exactly.

It is clear from Fig. 2.2 that there is a regime for which p_k is roughly constant, and that the extent of this regime tends to decrease with t. The values of ν predicted by the condition $\langle \tau \rangle_{\nu} + \sigma_{\tau,\nu} = t$ are 13 for $t = 0.8$ *sec* and 15 for $t = 2.0$ *sec* [slightly overestimated by Eq. (2.7), whose integer part yields 14 for $t = 0.8$ *sec* and 15 for $t = 2.0$ *sec*], in rough agreement with the values of about 12 and 14 respectively obtained from Fig. 2.2.

At this point a few important considerations are in order. The fraction $1 - p$ of molecules whose replication was incomplete will give rise to incomplete complementary single strands. Only when these incomplete replicas are close to completion will they be able to bind a primer in the next cycle, and thus be replicated. Therefore the efficiency p defined in Eq. (2.2) is an underestimation, since h_{N-1}, h_{N-2}, ..., h_{N-j} as well as qh_{N-1}, qh_{N-2}, ..., qh_{N-j} (for some $j < L_p$, where L_p is the length of the primers), will be part of the pool of templates in subsequent cycles. However, the dominant process will be the replication of the complete strand, which justifies the computation of p as in Eq. (2.2). There is another issue that needs some discussion. All the complementary strands arising from both complete and incomplete replication of a template can anneal to that template in subsequent cycles,

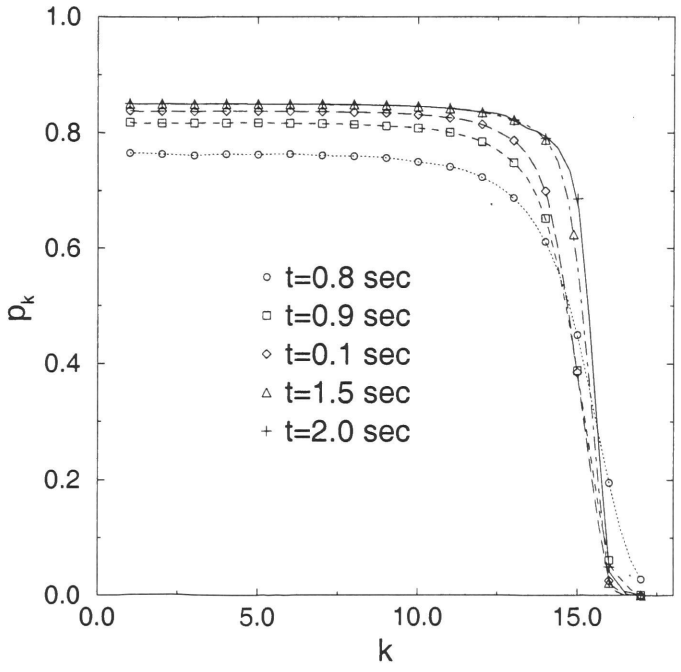

Figure 2.2: Efficiency p_k as a function of the cycle number k, for different polymerization times. The length of the template is $N = 45$; the other parameters are as in Fig. 2.1.

and therefore can act effectively as primers. Strictly speaking, at any given cycle $k \geq 1$ there will be a pool of primers of different lengths. An estimate of the concentration of "primers" arising from incomplete replication at cycle k is $\frac{1-p}{p}(1+p)^k[ss](0)$. This amount is always smaller than the concentration $(1+p)^k[ss](0)$ of completely replicated single strands which act also as potential "primers". As long as the concentration of incomplete replicas remain much smaller than the concentration of primers $[pr]$ Eqs. (2.1) will constitute a good approximation to the PCR process. Recall now that ν [see Eq. (2.7)] is equal or smaller than the number of cycles required for the concentration of primers $[pr]$ to match the concentration of completely replicated single stranded molecules. It follows that the approximation given by Eqs. (2.1) will break down only after the number of PCR cycles is bigger than ν and therefore our basic conclusions, contained in Eqs. (2.4)-(2.7), are not altered.

2.4 Statistical Analysis

As seen above, the efficiency p can be assumed to be constant for a number of cycles of PCR. The statistics of PCR can be readily computed under this assumption. The basic element in the analysis is the recursive relation that links the number of replicates after cycle number $n+1$, N_{n+1} in terms of N_n,

$$N_{n+1} = N_n + B(N_n; p) \tag{2.8}$$

where $B(N_n; p)$ is a random variable whose distribution is binomial with parameters N_n and p. The basis for this relation is that at the $n+1$-th cycle,

there will be not only the N_n molecules that were present at the previous cycle, but also the number of successful replication after N_n Bernoulli trials [Feller 68] each one with probability p of success. The number of molecules in the initial sample will be denoted by M_0.

The first moments of N_n can be easily computed from Eq. (2.8):

$$\mu_n \equiv \langle N_n \rangle = M_0(1+p)^n, \tag{2.9}$$

$$\sigma_n^2 \equiv \langle (N_n - \mu_n)^2 \rangle = M_0 \frac{1-p}{1+p}\left[(1+p)^{2n} - (1+p)^n\right]. \tag{2.10}$$

Furthermore, using the theory of branching processes [Harris 63, Athreya 72], a recursive relation between $P_n^{M_0}(k)$ (the probability that there are k molecules at cycle n, having started with M_0 of them) and $P_{n-1}^{M_0}(k)$ can be obtained

$$P_n^{M_0}(k) = \sum_{j=[\frac{k}{2}]}^{j_{max}} \binom{j}{k-j} p^{k-j}(1-p)^{2j-k} P_{n-1}^{M_0}(j), \tag{2.11}$$

(where $j_{max} = \min\{M_0 2^{n-1}, k\}$, and $[\frac{k}{2}]$ denotes the integer part of $\frac{k}{2}$), which when supplemented with the initial condition $P_0^{M_0}(k) = \delta_{k,M_0}$ allows us to compute $P_n^{M_0}(k)$ for any n. Figure 2.3 shows the form of these probability functions for $n = 10$ with $M_0 = 1$ in Fig. 2.3(a), and $M_0 = 50$ in Fig. 2.3(b), and different values of p. A remarkable resonance-like behavior can be observed in the curve corresponding to $p = 0.9$ and $M_0 = 1$ [wavy curve in Fig. 2.3(a)]. This phenomenon originates in the discrete nature of the process: if at the first cycle the system fails in replicating the only original

template, then the subsequent growth of the population will be as if there were nine cycles instead of ten. The other peaks correspond to the failure in replication in the first two cycles, three cycles, etc. This trait is characteristic of values of p between say 0.8 to 1. For smaller values of p the function looks smoother. A common feature of the curves in Fig. 2.3(a) is the existence of a power law regime in the region of small N_n, whose origin will be discussed later on. The behavior of the curves with $M_0 = 50$ is simpler: they are basically Gaussian curves, with a mean that increases with p and a variance that first increases and then decreases with p [see Eqs. (2.9) and (2.10)].

In order to understand the features described above, it is convenient to use the formalism of generating functions [Feller 68]. The generating function of $P_n^{M_0}(N_n)$ is simply $g_{n,M_0}(s) = \langle s^{N_n} \rangle$. Using Eq. (2.8) it is clear that $g_{1,1} = (1-p)s + ps^2$. It can be shown [Harris 63] that for a branching process

$$g_{n,M_0}(s) = g_{n-1,M_0}[g_{1,1}(s)] = \ldots = \left[g_{1,1}^{(n)}(s)\right]^{M_0}, \tag{2.12}$$

where we have denoted by $g_{1,1}^{(n)}(s)$ the *n-th* composition of $g_{1,1}(s)$ with itself, and used that $g_{0,M_0}(s) = s^{M_0}$ in the last equality. To proceed, we use the formalism of characteristic functions. The characteristic function $\phi_{n,M_0}(\omega)$ of the distribution of N_n having started with M_0 molecules, which is by definition the Fourier transform of $P_n^{M_0}(N_n)$ [Feller 68], is simply $\phi_{n,M_0}(\omega) = g_{n,M_0}(e^{i\omega})$. In terms of the characteristic functions, Eq. (2.12) implies that

$$\phi_{n,M_0}(\omega) = [\phi_{n,1}(\omega)]^{M_0}. \tag{2.13}$$

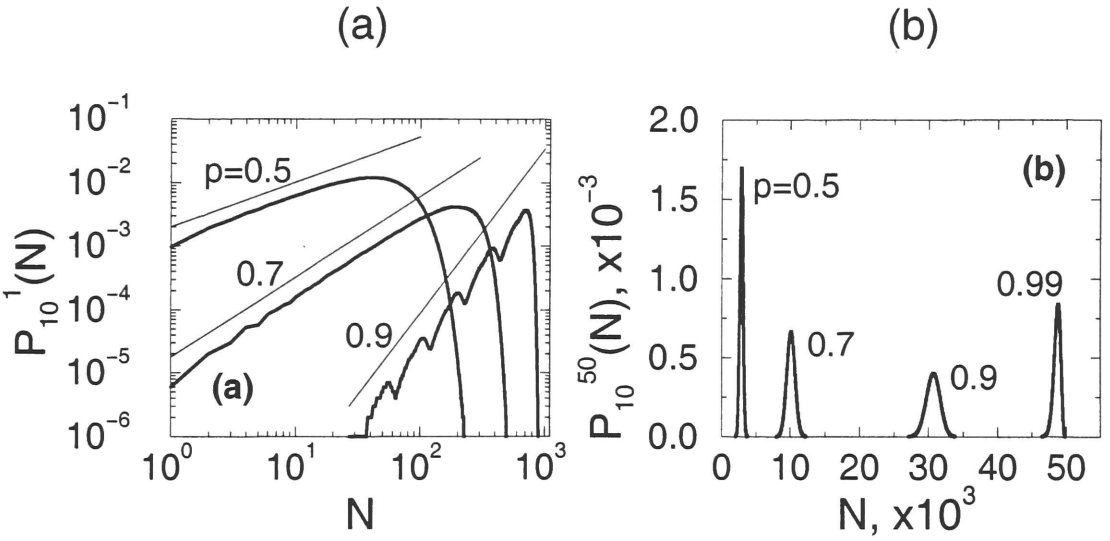

Figure 2.3: (a) *pdf* of the number of molecules after $n = 10$ cycles and $M_0 = 1$ of a branching process with constant efficiency p, in log-log scale. Notice the multi-modality for $p = 0.9$, and the power law regimes (straight lines). (b) Same as in (a) for $M_0 = 50$ (linear scale). The multi-modality has disappeared even for $p = 0.99$.

The characteristic function of the sum of M_0 *independent* random variables is simply the product of the characteristic functions of each of them. Therefore, the physical interpretation of the last equation is that the amplification cascades produced by each of the M_0 original molecules proceed independently, without interaction. From this observation and the central limit theorem it follows that as the number of molecules M_0 becomes larger, the distribution of N_n tends to a Gaussian. This explains the observed features of the *pdf*s of Fig. 2.3(b).

The behavior of the *pdf*s for finite M_0 in the limit of $n \to \infty$ is a little bit more interesting. In fact, it is clear from Eq. (2.13) that it suffices to study the case $M_0 = 1$, which we do next. We should stress that our study of the asymptotically large n regime does not aim at understanding the behavior of PCR when infinitely many cycles are performed. In fact we have shown in the previous Section that the efficiency can be considered as constant only for a finite number of cycles. Rather, the reason for studying this asymptotic regime is that the convergence of the finite n case to the $n \to \infty$ case is fast enough that many of the features arising for finite n are well explained by the study of asymptotically large n, most notably the power law behavior of the low N regime of Fig. 2.3(a). It follows from Eq. (2.12) that $g_{n,1}(s) = g_{1,1}[g_{n-1,1}(s)]$, which in terms of the characteristic functions and of the explicit expression for $g_{1,1}(s)$ becomes

$$\phi_{n,1}(\omega) = (1-p)\phi_{n-1,1}(\omega) + p[\phi_{n-1,1}(\omega)]^2. \tag{2.14}$$

Given that we are going to consider the limit of $n \to \infty$ and $\langle N_n \rangle = (1+p)^n$ [see Eq. (2.9)] diverges in this limit, it is convenient to use the random variable $\tilde{N}_n = N_n/\langle N_n \rangle$. Denote by $\theta_{n,1}(\omega)$ its characteristic function. It is easy to show that $\theta_{n,1}(\omega) = \phi_{n,1}(\frac{\omega}{(1+p)^n})$, which on using Eq. (2.14) yields

$$\theta_{n,1}(\omega) = (1-p)\theta_{n-1,1}(\frac{\omega}{1+p}) + p[\theta_{n-1,1}(\frac{\omega}{1+p})]^2. \tag{2.15}$$

Notice that Eq. (2.15) can be thought of as a dynamical system, that maps the point z_n to $z_{n+1} \equiv f(z_n) = (1-p)z_n + pz_n^2$. The function $\theta_{n,1}(\omega)$ $(-\infty \leq \omega \leq \infty)$ parameterizes a curve in the complex plane. In fact, the initial condition $M_0 = 1$ determines that $\theta_{0,1}(\omega) = e^{i\omega}$, which parameterizes the unit circle ζ_0. Subsequent applications of the map $f(z)$ to ζ_0 produces the new curves $\zeta_1, \zeta_2, \ldots$, which are parameterized respectively by $\theta_{1,1}(\omega), \theta_{2,1}(\omega), \ldots$. The study of the limiting behavior of the *pdf* of $\tilde{N}_n$ is thus associated with the study of the invariant curves of the map f. Notice that the map f has only two fixed points, one at $z = 0$ (stable) and one at $z = 1$ (unstable). Upon iteration, all the infinitesimally small straight lines with slope λ passing through the repelling point $z = 1$ will generate a curve C_λ which is invariant under f, that is $f(C_\lambda) = C_\lambda$. On the other hand, for any $z \neq 1$ such that $|z| \leq 1$, $|f(z)| < |z|$. Therefore the dynamics of this map brings all the points of ζ_0 (except for $z = 1$) to the origin. In the neighborhood of $z = 1$, ζ_0 is locally a straight line with slope $\lambda = \infty$, which upon evolution will become the invariant manifold C_∞. It follows that ζ_∞ coincides with C_∞, and $\theta_{\infty,1}$ parameterizes the invariant manifold of the map f, that crosses $z = 1$ parallel

to the imaginary axis. Figure 2.4(a) shows half the invariant manifold C_∞ corresponding to $p = 0.9$ (the other half is its complex conjugate), and on the same plot the imaginary part *vs* the real part of $\theta_{15,1}(\omega)$ (for positive ω). To the level of resolution of the figure no departures between the two curves are observed, meaning that the *pdf* of the number of molecules at 15 PCR cycles is well approximated by the limiting *pdf*.

This dynamical-system way of looking at the characteristic function of $\tilde{N}_n$ is very useful to understand the power law behavior of the *pdf*s of N_n. The argument goes as follows. Close to $z = 0$ (or equivalently, for large values of ω) the quadratic terms in Eq. (2.15) can be neglected, and the resulting approximate relation, $\theta_{\infty,1}(\omega) = (1-p)\theta_{\infty,1}(\omega/(1+p))$ accepts as a solution the ansatz $\theta_{\infty,1}(\omega) \approx A(\ln\omega)\omega^{\frac{\ln(1-p)}{\ln(1+p)}}$, where $A(x)$ is in principle any periodic function with period $\ln(1+p)$. The large ω behavior of the characteristic function $\theta_{\infty,1}(\omega)$ is then a power law, with logarithmically periodic modulations. That this is so is shown in Fig. 2.4(b), where we have plotted the absolute value of $\theta_{15,1}(\omega)$ for $p = 0.9$. The power law corresponding to the predicted scaling exponent of $\frac{\ln(1-p)}{\ln(1+p)}$ is shown as the straight line close to the curve in the log-log plots of Fig. 2.4(b). The implication of this results for the *pdf*s can be readily drawn. Recalling that the characteristic function and *pdf* are related through a Fourier transform, and that the Fourier transform of $|\omega|^\alpha$ (with an appropriate infrared cut-off) scales as $x^{-\alpha-1}$, we conclude that the *pdf* of $\tilde{N}_n$ should exhibit a scaling of the form $P^1_\infty(\tilde{N}_n) \sim \tilde{N}_n^{-\frac{\ln(1-p)}{\ln(1+p)}-1}$.

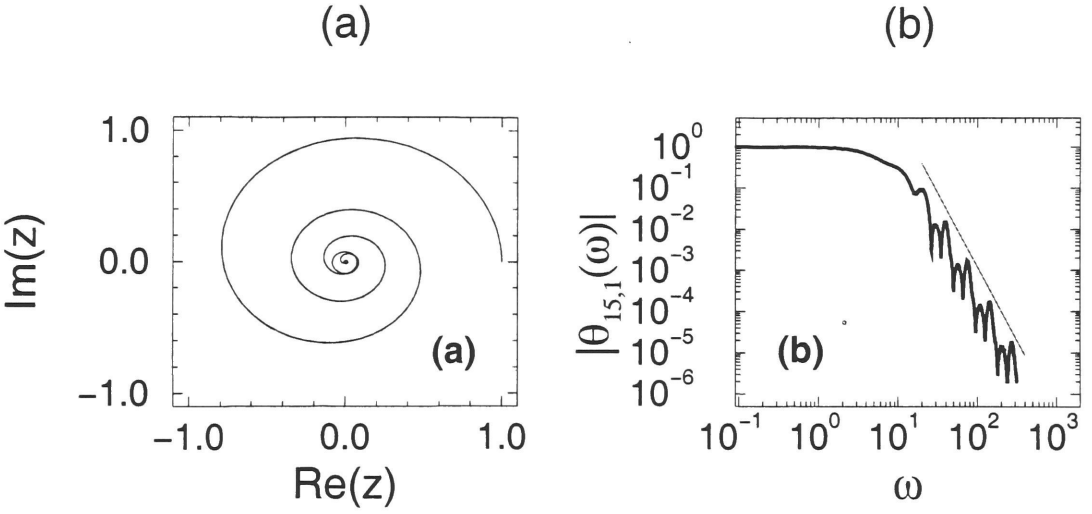

Figure 2.4: (a) The invariant manifold that crosses $z = 1$ tangent to the unit disk, of the map $z_{n+1} = (1-p)z_n + pz_n^2$ (with z in the complex plane), for $p = 0.9$. It is parameterized by $\theta_{\infty,1}(\omega)$. In the same plot the curve parameterized by $\theta_{15,1}(\omega)$ is shown, and cannot be resolved from the invariant manifold. (b) Absolute value of $\theta_{15,1}(\omega)$, and the predicted power law.

This scaling law is shown as the straight lines close to the curves plotted in log-log scale in Fig. 2.3(a).

In the following Section we apply some of the results presented so far to the problem of quantitative PCR.

2.5 Quantitative PCR

Although PCR is used mainly in a qualitative fashion, its potential for becoming an important tool in nucleic acid quantification in general [Ferre 92], and in medical research in particular [Clementi 93] has become clear in recent years. By quantitative PCR one means the use of the PCR to measure an unknown initial number of molecules M_0. A few techniques have been developed to that effect in the past, but the most widespread is probably the so-called competitive PCR (see, e.g., [Guilliland 90]). In this technique, the target, whose initial concentration is unknown, is amplified simultaneously with a standard, which is flanked by the same primers as the target and whose initial concentration is known. The standard should have a length different from that of the target, so that both can be resolved in an electrophoretic gel. The basic idea in competitive PCR is that *if the efficiencies of replication of the target and the standard are the same* then the ratio of the concentration of target to that of the standard is constant in the reaction. Measuring that ratio at cycle n (where presumably we have enough concentration to use densitometric measurements) we can solve for the initial concentration of target.

While this technique is very attractive, the basic assumption (the equality of the efficiencies in both species) has some drawbacks [Raeymakaers 93]. Basically, the potential problems arise in the dependence of the efficiency on the length of the DNA molecule. The longer molecule will experience a decrease in efficiency before the shorter one does, as predicted in Eq. (2.7). In any case the model presented here can be of use to assess the validity of the assumptions that go in the basics of competitive PCR.

In order for competitive PCR to work, the length of the standards must be within a narrow window: it has to be sufficiently different from the length of the target molecules to be resolved in a gel, and sufficiently similar to it in order for the equal efficiency assumption to work. The design of a good standard requires some ingenuity, and has to be done on a case by case basis. In what follows we will present a design for measuring M_0 without the need of a standard. Suppose we measure the concentration of a given DNA molecule after a number of PCR cycles on a sample whose M_0 is unknown. One might think that if we repeated the same measurement for a reasonable number of times (say around 100 times, given that PCR equipment with capacity for 96 vials is not uncommon), so as to measure the mean value and the variance of the concentration across that number of experiments, we would have two equations [Eqs. (2.9) and (2.10)] that can be solved for the two unknowns p and M_0. However, it can be shown that this procedure always yields two possible solutions for p and M_0, and there is no possible way *a*

priori, of choosing the right one. The reason for this is that for M_0 bigger than a few hundreds (which is nonetheless a small number of molecules), the distribution of N_n is Gaussian, and therefore determined only by the mean and the variance, which give the above mentioned ambiguous answer.

Consider instead the following scheme. We prepare two sets of samples S_1 and S_2, each with K identical preparations and whose initial concentration of a given double-stranded DNA molecule is unknown. We run (under conditions for which p can be considered approximately constant from cycle to cycle) n_1 cycles of PCR on set S_1, and n_2 cycles on set S_2, after which we measure the number of molecules in every sample. The averages ν_1 and ν_2 over the K preparations in S_1 and S_2, are estimates of the ensemble averages μ_{n_1} and μ_{n_2} corresponding to Eq. (2.9) for $n = n_1$ and $n = n_2$ respectively. We can use that formula to compute $m_0 = \nu_1^{-\frac{n_2}{n_1-n_2}} \nu_2^{\frac{n_1}{n_1-n_2}}$ as an estimate of the real M_0 and $\rho) = \nu_1^{\frac{1}{n_1-n_2}} \nu_2^{-\frac{1}{n_1-n_2}} - 1$ as an estimate of the real p. Of course these estimates make sense only if a measure of the error involved in the method is provided. It takes a simple calculation to show that,

$$\langle m_0 \rangle \approx M_0; \qquad \langle \rho \rangle \approx p, \tag{2.16}$$

and

$$\sigma_{m_0} \equiv \frac{\langle (m_0 - \langle m_0 \rangle)^2 \rangle}{\langle m_0 \rangle^2} \approx \frac{1}{M_0 K} \frac{1-p}{1+p} \frac{n_1^2 + n_2^2}{(n_1 - n_2)^2}, \tag{2.17}$$

$$\sigma_{\rho} \equiv \frac{\langle (\rho - \langle \rho \rangle)^2 \rangle}{\langle \rho \rangle^2} \approx \frac{1}{M_0 K} \frac{1-p^2}{p^2} \frac{2}{(n_1 - n_2)^2}. \tag{2.18}$$

In writing the last two equations we used Eq. (2.10). We tested these expressions in a set of very simple numerical simulations, whose details we are not going to report here except for saying that the PCR amplification was represented by the cascade given by Eq. (2.8). Under variations of all the parameters involved, Eq. (2.16)-(2.18) were in excellent agreement with the numerical results. To get a flavor of the precision of the method proposed, assume a simple example with $M_0 = 1000$, $p = 0.8$, $n_1 = 10$, $n_2 = 15$ and $K = 50$. Under these conditions the above equations predict that the estimate of M_0 will be correct within 0.5% (that is ± 5 molecules) and that of p will be correct within 0.1%! These estimates refer to the purely statistical errors, and they will be fairly small under typical conditions. In real experiments they have to be supplemented with the errors involved in the measurement of the concentrations. If M_0 and p fluctuated from sample to sample (due to inevitable differences in their preparations), the fact that we are averaging over K samples will screen these fluctuations. In this latter case, Eqs. (2.16) will still be in agreement with the average M_0 and p, and Eqs. (2.17) and (2.18), which can be easily generalized to include these fluctuations, will give their right order of magnitude.

2.6 Conclusions

We have presented a kinetic model for the PCR, which can be the basis for a more accurate application of quantitative techniques, as it provides a dy-

namical account of the probability of replication as a function of the physical parameters involved. These include the rate constants of the different reactions. Conversely, the model allows us to extract information on these rates from direct measurements of p. From a theoretical point of view, it can also be used in the description of *in vivo* and *in vitro* enzymatic polymerization processes [McAdams 95]. The statistical analysis of PCR under the assumption of constant replication probability shows new interesting phenomena. The scaling behavior of the *pdf* is an effect of the recursivity of the process, whereas the multi-modality is related to failures in replication during the first cycles. Although the latter is a phenomenon present only for a small number of initial molecules, it is not far from actual experimental conditions, and might be of relevance in a number of quantitative applications, including molecular computation.

Part V

Appendix

Appendix A

Gene expression in the brain: the voice of the crowd

We present here a method to map inducible nuclear proteins developed for the study of ZENK in Chapter 4, but having obvious general applications. The experimental methods specific to that chapter are presented in Appendix B.

A.1 A system to map inducible nuclear proteins

The time course of ZENK protein expression and its brain distribution after song stimulation have been recently determined using ZENK immunocytochemistry (ICC) [Mello 95]. Due to its sensitivity, high spatial resolution and good tissue preservation, ZENK ICC can be counted among the best methods for functional mapping of brain areas involved in natural behaviors in vertebrates. Although other systems for tissue mapping are available

[Alvarez-Buylla 88, Hibbard 96], we decided to develop an automated system for ICC analysis that takes full advantage of the intrinsic properties of the ZENK ICC methodology. Our system accurately maps the spatial distribution and labeling intensities of immunopositive cell nuclei within a region of interest. Applied to songbirds, it allows for a high-resolution description of the brain's ZENK response to song stimulation ([Ribeiro 98], discussed in previous chapter), assaying the physiological state of very large neuronal ensembles.

We will describe an automated system that accurately maps tissue sections stained by immunocytochemistry for an inducible nuclear protein. The sections are scanned with a computer-controlled microscope setup hooked to a CCD camera. Raw images captured at high resolution are filtered using highly selective criteria for the recognition of labeled cell nuclei. The total population of recognized labeled nuclei is then divided into separate bins, according to their labeling intensities. Finally, information about both the position and labeling intensity of labeled nuclei is represented in average density maps. The system was optimized for the quantitative mapping of neuronal cells expressing the inducible gene ZENK in the brain of songbirds, in response to stimulation with song, but should be of general applicability for the mapping of inducible nuclear proteins [Cecchi 98].

A.2 Method

A.2.1 Tissue preparation

We analyzed brain sections from adult female canaries (*Serinus canaria*) stimulated with playbacks of a variety of auditory stimuli, including a whole canary song. Birds that heard only silence constituted an unstimulated control group. Data from about 110 animals were used at various stages in the development of this mapping system. The animals were stimulated for a period of 30 min., essentially as described elsewhere [Mello 95, Ribeiro 98]. The birds were killed and perfused with fixative 60 minutes after the end of stimulation, which is well within the time of peak ZENK protein expression [Mello 95]. The brains were dissected, frozen and sectioned (20 μm) in the parasagittal plane (Fig. A.1) and mounted onto salynated slides. Particular care was taken to ensure that all brains were cut in the same orientation.

A.2.2 ZENK ICC

We used an anti-ZENK polyclonal antiserum (C-19; Santa Cruz Biotech., Santa Cruz, CA, USA) and a previously described ICC protocol [Mello 95]. For detection of antigen-bound primary antibody we used the avidin-biotin-peroxidase method (ABC Elite kit, Vector Laboratories); the final cell labeling was the result of 3,3'-diaminobenzidine (DAB) precipitation over cell nuclei, with nickel intensification. All sections to be compared were processed in parallel, using the same batches of reagents.

A.3 Imaging

A.3.1 System setup

Sections are scanned using a computer-controlled microscope setup, consisting of:

- Silicon Graphics *Indigo*^{2} Workstation running Unix, with

  - 195 MHZ IP28 processor
  - MIPS R10000 processor chip, revision 2.5
  - MIPS R10010 floating point chip, revision 0.0
  - 32 Kbytes data cache
  - 32 Kbytes instruction cache
  - 1 Mbyte secondary unified instruction/data cache
  - 320 Mbytes main memory
  - SOLID IMPACT graphics board
  - IMPACT compression board, unit 0, revision 0:0
  - Software: the code of the navigator program (see below), as well as the analysis, was written in C language.

- Olympus BX60 upright microscope, with

- UPLFL UPlan Fluorite 40X/0.75 and 4X/0.13 objectives
- Video Adapter with 0.3X reduction lens
- CCD camera Sony XC-77CE PAL-N (768 x 576 pixels)
- Motorized Stepper Stage, 4x3 inches travel for Olympus BX, Prior Scientific model 500-H101
- Microstepping Motor Control Processor, 3 axis, $1\mu m$ precision, Prior Scientific model 500-H128V3, controllable from RS232 port.

A.3.2 Description of Pasting Algorithm

The "navigator" program (Fig. A.2) scans at low resolution (4X objective) the entire slide containing the sections to be analyzed, pasting the frames with the position obtained from the stage. On screen, polygons are drawn surrounding the sections of interest, as well as the area of the slide in which the average background illumination over the glass will be computed. This area is selected in each slide from a region of the coverslip without tissue (Fig. A.2), and 400 frames are averaged. To compensate for inhomogeneities in illumination, each frame is divided pixel-wise by the average background.

Although information about xy position is provided by the stage, it is reliable only up to 1 μm. With the 40X objective, the final resolution of the system is 3 pixels/μm, which means that a typical cellular nucleus spans around 15 pixels in diameter. Therefore, a precision of 1 μm in 200 μm of

scanned tissue is not enough to reliably paste hundreds of frames without distorting the circular shape of a nucleus. We therefore developed a pasting algorithm that achieves sub-micron precision, which will be described below.

In the first step, the system scans the tissue, acquires frames through the CCD camera, and stores them as a JPEG compressed movie file [Iris Prog.], together with the position returned by the stage, until the entire region of interest is covered. The overlap between consecutive frames is 150 pixels in x (out of 768) and 100 in y (out of 576). In the case of NCM, this results in approximately 300 frames, with a total size of around 150 MBytes. For each slide, the focal plane of the initial frame is determined manually. Autofocus is then implemented in the consecutive frames, in the following way: a focus score is computed as the absolute value of the gradient of the intensity integrated on a 127x127 central box; this is done for 10 focal planes around the current plane, in steps of 10 μm; the plane that maximizes the focus score is selected.

The next step is performed off-line. Each frame is processed for object recognition (cell nuclei, tissue irregularities, anatomical landmarks) using an algorithm previously described [Bourdieu 95]. The relative position of all objects recognized in each frame is used to refine the absolute position of the frame. This is implemented as follows:

1) The first frame is selected as the origin of the coordinate system.

2) With the information provided by the stage, the next overlapping frame

is selected. A histogram of inter-object distances between the frames is computed: using the relative position of the objects in each frame (equivalently, superimposing the frames), the vectorial distance is computed between each object in one frame and every object in the other frame; the resulting x, y values are then added to the x- and y- histograms.

2) The most populated bins in x and y are chosen, so that the corresponding values form the distance vector between the frames. This vector is used to correct the inter-frame distance provided by the stage (Fig. A.3).

3) These inter-frame distance vectors are used to re-constitute the final image of the entire area of interest, iterating the process for all successive frames . We have estimated that the reconstruction error of this method is less than 1 μm in 500 μm of scanned tissue.

After this off-line processing, the final image is stored as a JPEG compressed movie file, with 1008 x 1008 pixels frames. To avoid discrepancies on the borders, there is an overlap between consecutive frames of 100 pixels. For a typical NCM (1 mm x 1.5 mm), the final movie file is around 8 MBytes in size, and the whole process takes about 10 minutes.

A.3.3 Description of Recognition Algorithm

For each pixel we measure the optical density (O.D.) resulting from DAB-nickel precipitation. To implement an edge detector, we first approximate the "laplacian" of the image by multiplying each pixel value by σ/z, where σ and z are, respectively, the variance and mean intensity on a 3x3 kernel centered

at each pixel. The image thus obtained has peaks where the variation of the intensity relative to the local mean (in the scale of 1 μm) is high, and valleys where the relative variation is low. Therefore, it works as an effective detector of edges whose length scale is in the order of 1 μm, compatible with the definition of the nucleus border under the microscope. The resulting images are then processed for object recognition, based on an algorithm previously described [Bourdieu 95]. The support S of each object is computed filling the contours produced by the "laplacian" (typically "rings" for labeled nuclei). Anything outside the objects' support is considered as tissue, on which the average (non-specific) background labeling is computed. For each recognized object, we compute:

- The ellipse E with constant intensity that best fits the shape and intensity distribution on the support. This is done by diagonalizing the inertia matrix,

$$N_{ij} = \sum_{k \in S} i(k)j(k)I(k) - < i >< j > M \tag{A.1}$$

where $i(k), j(k)$ is the x,y position of pixel k, $I(k)$ its intensity, and

$$< i >= \sum_{k \in S} i(k)I(k)/M \tag{A.2}$$

$$< j >= \sum_{k \in S} j(k)I(k)/M \tag{A.3}$$

$$M = \sum_{k \in S} I(k) \tag{A.4}$$

- The eccentricity, as the ratio of the radia of E (eigenvalues of N_{ij}),

$$E_C = R_{min}/R_{max}$$

- The variance of the actual contour respect to the contour of E,

$$V_C = \left[\sum_{k \in S} \phi(k) + \sum_{t \in E} \psi(t)\right] / MT \tag{A.5}$$

where $\phi(k) = 1$ if k is outside E, and 0 otherwise; $\psi(t) = 1$ if t is outside S, 0 otherwise; and MT is the total number of pixels in S.

- The variance of the intensity in S,

$$V_I = \left[\sum_{k \in S} I(k)^2/MT - (M/MT)^2\right] / M \tag{A.6}$$

- The labeling intensity relative to the average background labeling B,

$$I_A = M/(MT \cdot B)$$

These measurements are then used to filter the objects:

$$3\mu m \leq diameter \leq 10\mu m;$$

$$E_C \geq 0.3;$$

$$V_C \leq 0.5;$$

$$V_I \leq 1.5;$$

$$I_A \geq 1.4;$$

only objects satisfying all of these requirements are considered labeled cell nuclei. Finally, the position and labeling intensity I_A of all labeled nuclei are stored for further processing. An example of the application of this procedure to an NCM section is shown in Fig. A.4.

A.3.4 Determining the outline of the region of interest

On the screen, a perimeter is drawn over the raw reconstituted image, so as to include the entire region of interest. In our particular case, the rostral boundary of NCM was a straight line along the major axis of field L2a. The dorsal, ventral and caudal boundaries of NCM were naturally given by the ventricular zone. An outline is then defined as the smallest convex polygon (convex hull) that includes all objects recognized as objects in the tissue within the drawn perimeter [Cormen 94]. An algorithm based on this definition is implemented to compute the outline of each section [1]. This procedure works very well for NCM because it has a convex shape; the generalization to other brain areas with convoluted boundaries like the cerebellum is certainly not straightforward, but we think it could be based on the same idea.

A.3.5 Definition of a standardized outline of the region of interest

To define a standardized outline which can be used as an anatomical reference for comparison across maps, we:

[1] The code to compute the convex hull of a set of points was written by Ken Clarkson, and can be down-loaded from *http://netlib.bell-labs.com/netlib/voronoi/*

1) Determine the outlines of the region of interest from a large number of animals, ideally the total population to be included in a particular study or comparison. Each individual outline is calculated as described in **2.3.4**.

2) Align the individual outlines relative to each other using a Monte Carlo algorithm [Press 94] that minimizes local misalignments and separation of the outlines. The energy function penalizes the distance d_{nn} and cross product (maximal for orthogonal segments) of nearest neighbor segments of the matching convex hulls, and integrates these measures along the contour. The cross product, being proportional to the length of the segments, also weighs the extent of the misalignment. A term measuring the variance of the distance between nearest neighbor segments is also included in the energy function, whose precise expression is:

$$\mathcal{E} = \left[0.01\langle d_{nn}\rangle\sqrt{\langle d_{nn}^2\rangle - \langle d_{nn}\rangle^2}\right] \sum \|\mathbf{v}_1 \times \mathbf{v}_2\|$$

where $\langle\rangle$ is the average along the contour, and $\mathbf{v}_1, \mathbf{v}_2$ the nearest neighbor segments in each contour.

3) Compute the standardized outline as the average of all outlines.

A.3.6 Scaling of maps

Individual maps are automatically aligned and scaled with respect to the standardized outline. For this, each map is:

1) Uniformly scaled so that its area matches that of the standardized outline.

2) Rotated and moved along the horizontal and vertical axis so as to align its orientation to the standardized outline, using the alignment algorithm described in **2.3.5**.

3) Scaled independently along x and y to minimize the variance of its outline respect to the standardized outline (Fig. A.6). The variance is measured computing the distance between nearest neighbor segments of the outline and the standardized outline.

A.4 Representation of the data

A.4.1 Classification of cells according to labeling intensities

In order to simultaneously represent the spatial distribution and the labeling intensity of recognized cells, we classify the cells into discrete bins according to their labeling intensities. In order to obtain a normalized classification so that maps can be compared, the total population of cells in a particular study should be included. We find that a classification into three bins (low, medium and high labeling intensities) is a convenient compromise between content and simplicity of presentation. The bins are defined so that each contains a third of the total cell population, i.e., a third of all the cells mapped, across sections. Each bin is then assigned a color; blue, green and red correspond to low, medium and high labeling.

A.4.2 Computation of stimulus-elicited averaging maps

The scaling and alignment of the sections with respect to the standardized outline is performed independently for each one, and only then the four maps corresponding to each stimulus group are averaged into a single response map. To generate average density maps, the set of points within each map is convolved with a gaussian kernel of 50 μm radius and the maps from all animals presented with the same stimulus are averaged. The three bins of labeling intensity are processed in parallel, so that three density maps - one for each bin - are independently generated per animal and finally superposed. The density of labeled cells (number of cells/μm^2) within each bin is represented by color brightness. The superposition of color-coded maps results in mixed populations of cells, which are represented in the average density maps by a myriad of hues derived from combinations of the blue, green and red bins (Fig. A.7B and C).

For each particular stimulus in our study (a total of 20), one section from each of the four different animals was mapped and included in the average map. The only criterion to select a section was its position relative to the mid-line (1100 μm to 1300 μm), so that all sections were comparable across animals. We decided to apply a lower intensity cutoff, so that only cells with labeling above 1.4 times the background labeling were considered. This strategy proved to be satisfactory for the discrimination of the stimulus-elicited maps, however, we will discuss below in Section 3 the possibility of lowering

the cutoff while preserving a reasonable reliability in nuclei recognition.

A.5 Testing the reliability of the system

To assess the reliability of the recognition procedure, we tested it in comparison with trained experimentalists. A selected test region of a NCM section was scanned by the system to recognize cells labeled above the background. On the same region, an exhaustive search of labeled cells was performed, using a consensus criterion between two subjects. The result of this search was then compared to the automatic procedure.

To test the reliability of the scaling and alignment procedure, we computed the ratios of the areas defined by individual NCM maps to that of the standardized NCM outline. In addition, we superimposed the outlines defined by all the individual maps used for the determination of the standardized NCM outline.

A.6 Results

The results reported here were obtained from the group of 80 animals used to develop the system.

To perform the test of reliability of nuclei recognition, we selected an area of approximately 450 μm x 900 μm in a section of an animal stimulated with song. The system identified 355 nuclei with labeling above background in that region. The comparison with the exhaustive search is as follows:

- In 10 cases, the system mis-recognized two closely apposed cells, connected by an isthmus of high labeling, as a unique cell.

- The system failed to recognize 96 cells (false negatives); all of these false negatives represent cells with low labeling intensity, as can be appreciated in Fig. A.5. The reliability within each bin is high for relative labeling above 1.2 (86%) and increases to almost 99% for intensities above 1.4. This is consistent with the 20% (STD) variability of the background labeling that we have observed across all sections, i.e., the choice of a unique background labeling per section precludes the recognition of weak cells.

As mentioned in Section 2.4.2, for our particular study we set the lower cutoff for cell labeling intensity at 1.4 above background, which represents a reliability of 99%; nevertheless, the test shows that the cutoff could be confidently lowered to 1.2.

The efficacy of the scaling and alignment procedure can be appreciated in Fig. A.6. Panel A shows how close individual outlines are aligned relative to the standardized outline after the procedure. Panel B shows that the areas of individual outlines approximates that of the standardized outline.

The total population of labeled cells observed in the 80 animals was 36,000. The ranges of labeling intensity to equally populate the three intensity bins were 1.40 to 1.62 (blue), 1.62 to 1.77 (green), and 1.77 to 2.68 (red). Fig. A.7C depicts how the maps of the three bins of labeling intensity

were superimposed so as to generate an average density map of the 4 animals stimulated with a whole song. In this final map, therefore, numbers of labeled nuclei are represented within three ranges of activation. It should be stressed that density maps do not represent pixels or densitometric unit per area, but rather absolute numbers of labeled nuclei per unit area. Note that for the particular stimulus in Fig. A.7C, a whole song, the three intensity populations are not balanced, because the song elicits a strong response, resulting in a predominance of the red bin (high expression level). As detailed elsewhere [Ribeiro 98], a classification of the cell nuclei population into three discrete bins of labeling intensity provides essential information for the discrimination of ZENK expression patterns resulting from different auditory stimuli.

Finally, the interanimal variability in terms of cell labeling intensities was small within each stimulus group. This is shown in Fig. A.7A, which depicts the distribution of labeling intensities for the group of animals stimulated with a whole song. The same is also true for several other auditory stimuli, as reported elsewhere [Ribeiro 98]. Thus, different auditory stimuli elicit distinct labeling distributions, reflecting the specificity of NCM responsiveness; in consequence, the different stimuli can be rather easily discriminated by their characteristic profiles.

A.7 Discussion

We have described here a procedure that allows us to perform automated recognition and quantification of labeled nuclei using ICC. We would like to discuss below two separate classes of problems that we had to deal with: quantification of ICC labeling and automated recognition of labeled nuclei.

A.7.1 Quantification of ICC labeling

We used an ICC protocol based on the deposition of DAB, a chromophore that provides stable staining under bright field illumination, and can therefore be examined repeatedly over extended periods of time [Kugler 90]. Due to the non-linearity of DAB deposition, one cannot directly infer the precise protein content of labeled nuclei. For classifying nuclei into discrete bins of labeling intensities, however, the only assumption needed is that the staining levels be monotonic with the amount of protein. Alternatively, fluorescent detection systems could minimize the non-linearity issue, but present their intrinsic disadvantages, particularly the short-lived light emission after ultraviolet excitation. Irrespective of the detection system employed, quantification of ICC labeling is not a trivial task, and a substantial amount of work is required to solve the issue of variability. We have identified and addressed four main variability sources, which we discuss below in detail.

A.7.2 Labeling variability

It is well known that local differences in tissue perfusion and fixation, as well as access to reagents during the ICC procedure, can have a strong influence on ICC staining levels. It is therefore essential to minimize these factors by reacting all sections in parallel, with the same batch of reagents. Any remaining differences have to be dealt with by using a normalization procedure. We have previously described that background labeling in our system is independent of specific cell labeling, as it persists when the primary antibody is pre-absorbed with its specific antigen, or even when incubation with the primary antibody is omitted from the ICC procedure [Mello 95]. We have therefore normalized cell labeling values to that of the background labeling. An alternative procedure commonly adopted to address labeling variability consists of normalizing labeling respective to a different brain area. This has the drawback of potentially accentuating variations in staining due to different access to perfused fixative, different position in the reaction vial, and differences in the response of different brain areas to the experimental condition (in our case song stimulation). Our mapping procedure allows us to compute the precise value of background labeling, and therefore to normalize labeling to the background within the region mapped. To our minds, this is a more rigorous way of dealing with the issue of labeling variability.

A.7.3 Human error in the identification of labeled cells

A major concern we had at the beginning of our work was the unreliability of humans to map a very large number of cells, over long periods of time (weeks to months), while using a consistent set of criteria. Our computerized system applies a single set of criteria for cell recognition to all maps, allowing for the comparison of animals presented with different stimuli, or mapped at different times.

A.7.4 Histological variability

Even when extreme care is taken to cut the brains from different individuals in the same anatomical plane, it is inevitable that some variability results in section size, shape, and orientation. Any attempt to compare sections across different animals must deal with this issue. Our scaling and alignment procedure in reference to a standardized outline of the region of interest allows us to normalize the entire data set in terms of size, shape and orientation of sections. Notice that the algorithm used for alignment and scaling does not depend on the ICC labeling patterns, but rather takes into account only the anatomical boundaries of the region of interest; this precludes the introduction of bias in the procedure, due to differential tissue responsiveness.

A.7.5 Intrinsic interanimal variability

After minimizing or eliminating the main extrinsic sources of variability, one still has to face the inter-individual variability of responses within the region of interest. We dealt with this issue by averaging the data of all animals per stimulus group. As the experimental procedure requires that the animal be killed, there is a limitation in the number of stimuli that can be reasonably compared in a single study. This sets the limits of applicability of the ZENK mapping procedure for physiological studies, since a minimum number of animals must be used per condition or stimulus to guarantee a meaningful interpretation. In the case of auditory responses in NCM, including 4 animals per group is sufficient to grant statistical confidence to the differences observed across groups [Ribeiro 98].

A.7.6 Automated detection of labeled nuclei

The automated detection of ICC labeled nuclei in our system is an iterative process. For a first and rough detection of labeled objects in the tissue, we employ a low cutoff for the "laplacian", or gradient of labeling intensity, so that labeled nuclei as well as similarly stained debris are included. This is more efficient than using directly the value of labeling intensity, as it eliminates artifactual objects with high labeling levels but fuzzy boundaries. Next, a set of restrictive filters is used to refine the initial selection. Finally, a more stringent cutoff for labeling intensities is applied, that takes into account

the background staining variance. It should be pointed out that no single criterion is enough for a reasonable recognition of labeled cells; when applied in conjunction, however, the series of filters is highly efficient. Interestingly, the development of the system itself consisted of an initial educated guess of what should be the different filters and their values; with successive trials, the filter values became gradually more stringent.

Our automated system reliably recognizes 80 % to 90 % of the cells identified by trained human observers which have labeling intensities 20% above the background. Even though tissue artifacts affect the recognition to a minor extent, the main source of discrepancies are weakly labeled cells. Human observers have particular difficulty in consistently identifying cells with low labeling and deciding how they compare to tissue background; this issue is easily solved by the automated system. On the other hand, the system has some difficulties in identifying weakly labeled cells where the labeling is spatially restricted to a small domain within the nucleus, or in separating closely opposed cells. The net result of the discrepancies between the human observers and the automated mapping system is an underestimation of the number of labeled nuclei by the latter. There are, however, considerable advantages of the automated system that constitute a satisfactory trade-off, including:

- criteria consistency,
- accurate spatial localization,

- quantification of labeling intensities, and
- the ability to quickly process massive amounts of data.

Attesting to the usefulness of the system, the information it generated for the mapping of the ZENK responses in NCM to various auditory stimuli was enough to provide for a full classification of the resulting ZENK expression patterns [Ribeiro 98].

Regarding the applicability of this method to areas with a more complex anatomy, we want to discuss two different aspects. First, as mentioned before, NCM is a convex area, which helped in the computation of the boundary, and subsequently in the comparison and averaging of sections from different animals. Convoluted boundaries, like the cerebellum's, cannot be naturally approximated by the convex hull of the tissue. There are, nevertheless, formal extensions of the convex hull, like the α-shape [Edelsbrunner 83], which could be implemented in dealing with non-convex outlines. The second aspect is the reconstruction of the three-dimensional anatomy of an entire brain area of a single animal, and eventually the comparison of this 3-D structure across animals. This is a more difficult task, given the lack of guidelines as to how successive sections should be stacked. For particular brain areas, the problem could be attacked on the basis of prior knowledge on the anatomy. This problem notwithstanding, our method can provide the basis for any attempt towards the reconstruction of entire neuronal networks undergoing genomic activity in the brain.

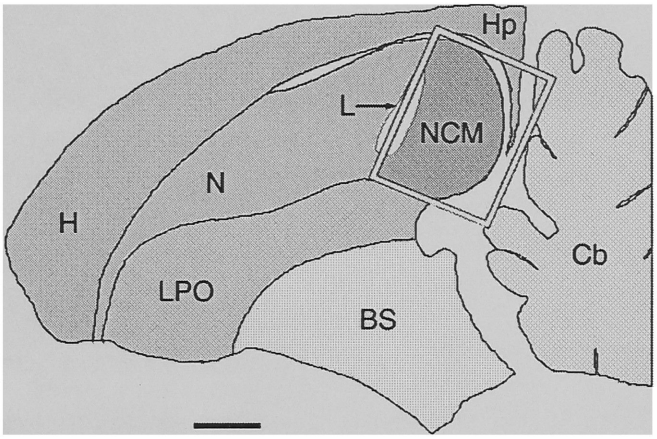

Figure A.1: Parasagittal section of the canary brain at the level of NCM (1100 to 1300 μm lateral to mid-line). Abbreviations: BS, brainstem; Cb, cerebellum; H, hyperstriatum; Hp, hippocampus; L, subfield L2a of field L; LPO, lobus parolfactorius; N, neostriatum; NCM, caudomedial neostriatum. Scale 1 mm.

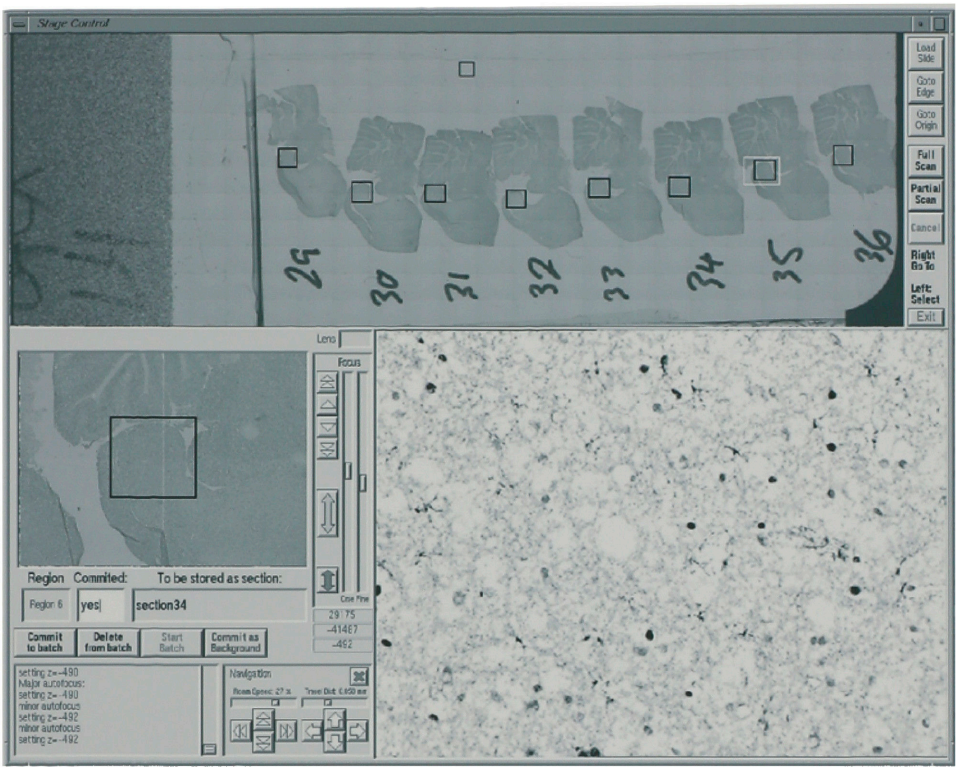

Figure A.2: "Navigator" interface. The upper panel shows the scan at low resolution of the entire slide. The lower right panel shows a single frame at high resolution. The lower left panel shows the selection of the area of interest (NCM). The black rectangles represent the regions selected for scanning, the topmost rectangle encloses the region selected for the computation of background illumination, and the white rectangle indicates the current position of the objective.

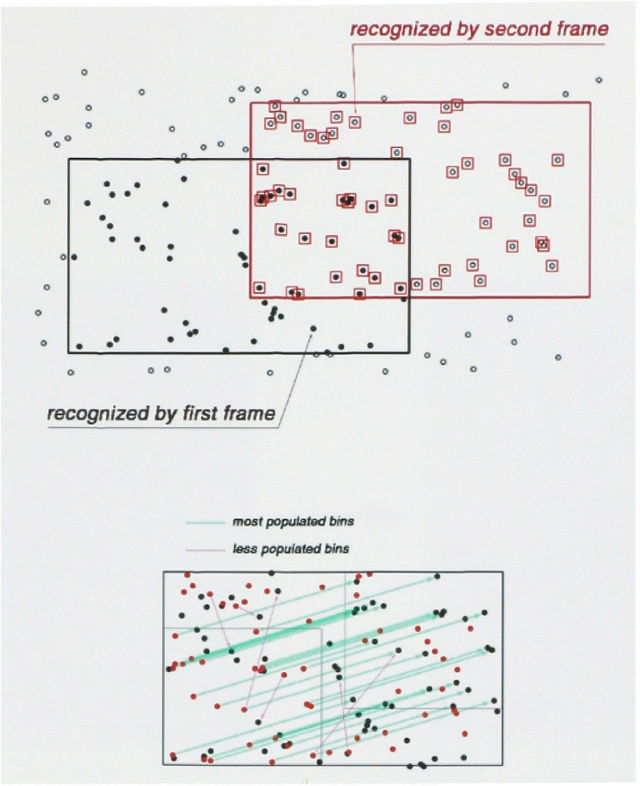

Figure A.3: Refinement of the position of the frames. The object detection algorithm provides the position of the objects found in each frame, relative to it. The frame position provided by the stage permits the identification of overlapping frames (Upper panel). When superimposed, the overlapping frames share many objects, but separated by a distance equal to the separation between frames. To compute an xy-histogram of inter-object distances, the vectorial distance between an object in the first frame and all objects in the second is calculated for each object of the first frame. These distances populate the xy-histogram randomly, except for the shared objects, resulting in that the most populated bin corresponds to the actual frame separation (Lower panel).

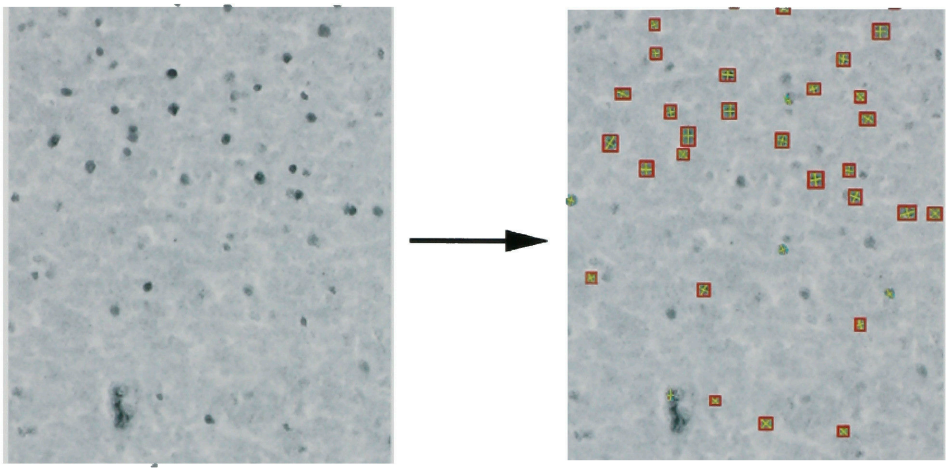

Figure A.4: Automated recognition of labeled nuclei. The left panel depicts a region of NCM after ZENK ICC, where several labeled objects can be observed. The right panel shows which of these objects are recognized as nuclei by the system. Recognized nuclei are marked by a yellow cross and surrounded by a red box. Notice that some objects are marked only by the yellow cross, indicating that not all the criteria for nuclei recognition are met by these objects.

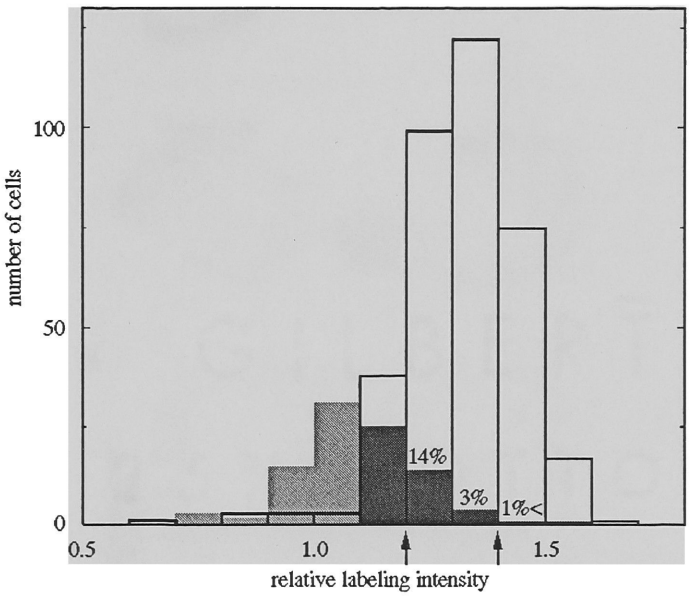

Figure A.5: Number of cells per intensity bin recognized by the system (hollow bars) and false negatives relative to exhaustive search by human observers (light shade). The dark shade indicates false negatives with intensities above 1.2 relative to background. The number of false negatives is less than 7% of the total number of cells with intensity above 1.2, although it is 14% relative to the first bin and decreases very fast for higher intensities. The arrows indicate the chosen intensity cutoff (right) and the lowest reliable cutoff (left).

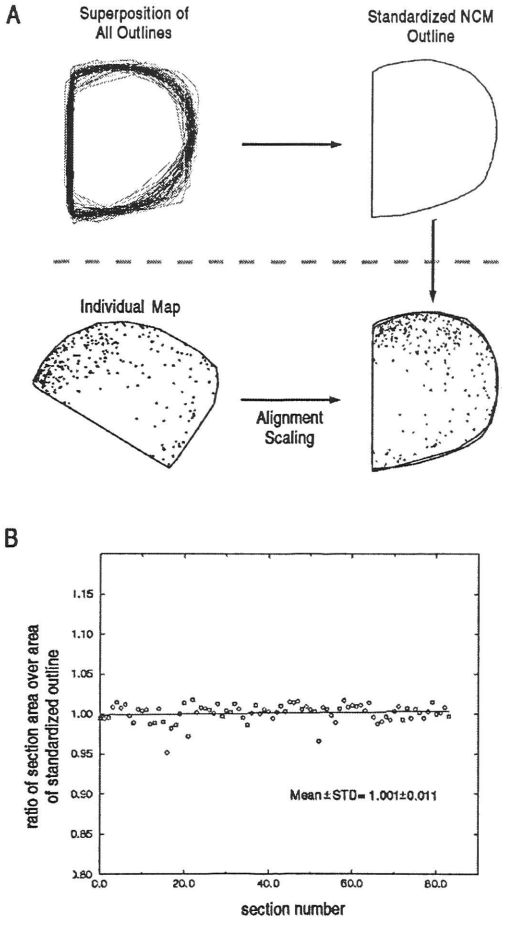

Figure A.6: Panel A: Scheme of alignment and scaling algorithm. The top part shows how the standardized NCM outline is extracted from the superposition of several (80) NCM outlines, after their alignment and scaling. The bottom part shows how individual maps are aligned and scaled respect to the standardized NCM outline. Panel B: Areas of individual outlines after alignment and scaling are plotted relative to the area of the standardized NCM outline.

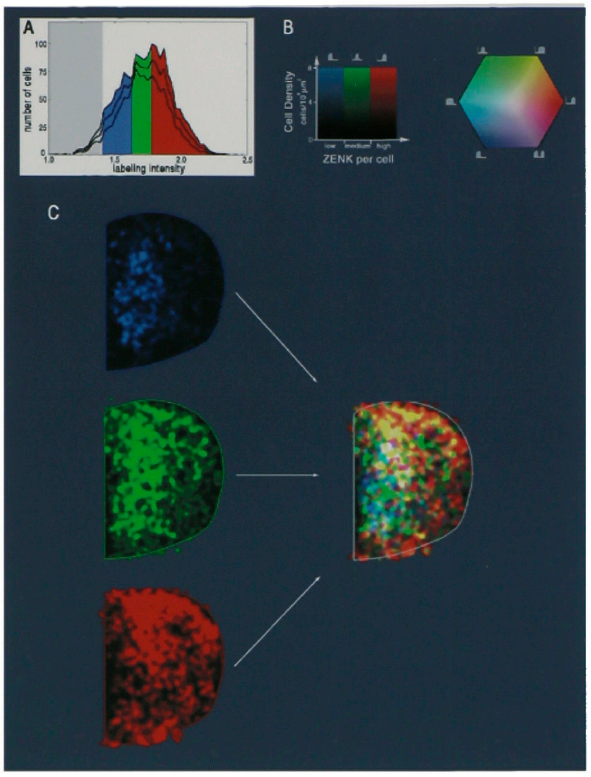

Figure A.7: Panel A: Distribution of ZENK-labeled cells resulting from presentation of a whole song. The central line indicates mean values of the number of cells with a given labeling intensity, while the upper and lower lines represent the variance ($\pm$ sem). The shaded area indicates the lower cutoff adopted; the colored areas indicate the color-coding ranges according to three bins of labeling intensities (blue, 1.4 to 1.62; green, 1.62 to 1.77; red, 1.77 to 2.68; for the determination of these values, see Results). Panel B: The key on the left shows how brightness in each of the three color-coded ranges reflects cell density; the key on the right indicates how mixed cell populations result in various color mixtures. Panel C: Map of ZENK expression in NCM of canary resulting from stimulation with a whole song. The separate density maps for each of the three labeling intensity bins are shown on the left. The map resulting from the superposition of the three partial maps is shown on the right.

Appendix B

Experimental Procedures for Chapter 4

B.1 Animal groups

Our study consisted of 20 groups (n=4 per group) as follows: 17 different stimuli, a silence control, plus two groups in the habituation experiment, for a total of 80 adult canaries (Serinus canaria). All animals were obtained from the aviaries of the Rockefeller University Field Center, Millbrook, NY. We used females, which do not sing under our experimental conditions, to avoid ZENK induction by auditory feedback during singing.

B.2 Stimulation

Each bird was acoustically isolated for a day to minimize basal ZENK expression, and then presented with a particular stimulus. The whole song and all natural syllables were obtained from canary repertoire. Synthetic

whistles are pure sinusoidal tones whose amplitude envelopes were smoothed to resemble the elliptic shape of natural whistle envelopes. Guitars are digitally sampled guitar notes from which harmonics were filtered down to 30 dB below the intensity of the dominant frequency; the resulting intensity of harmonics during presentation of guitar notes was therefore 10 dB below ambient noise levels (see below). The song studied here (Figure A.1A) has a duration of 10 seconds. The average duration of other stimuli is as follows: 365 ms for natural whistles, chords, synthetic whistles and guitar notes; 720 ms for combinations; and 25-50 ms for modulations. These stimuli were presented in quick repetition for 10 s. to match the duration of the whole song; the intervals between consecutive renditions were 10-20, 25 and 50-100 ms, respectively, for the three groups above. The resulting rates of presentation were approximately those at which the different natural syllables occurred in the canary song studied here; artificial stimuli were presented so as to match the protocol for natural whistles. Stimulation was delivered at an average intensity of 70 dB sound pressure level (SPL), in 60 such blocks of 10 seconds, equally spaced by 20 seconds of silence (total duration of 30 min). Unstimulated controls were kept in silence (average ambient noise intensity of 50 dB SPL). In the habituation experiments, the birds were initially habituated to two individual whistles presented separately , in alternating blocks of 10 s each (separated by 20 s of silence) for 7 hr. One group was then presented with the same two whistles combined as a sequence for 30 min (also in a

10/20 s schedule), while the control group continued to be presented with the same two whistles separately.

B.3 Immunocytochemistry

Animals were killed and perfused with fixative 60 minutes after end of stimulation, which is well within the time of peak ZENK protein expression [Mello 98a]. The brains were dissected, frozen and sectioned (20 ?m) in the parasagittal plane. Particular care was taken to ensure that all brains were cut in the same plane. We used an anti-ZENK polyclonal antiserum (C-19, Santa Cruz Biotech., Santa Cruz, CA, USA) and a previously described ICC protocol [Mello 98a]. Cell labeling was the result of 3,3'-diaminobenzidine (DAB)/nickel precipitation over cell nuclei, where ZENK protein is localized. Staining variability was minimized by reacting all sections in a single batch and normalizing all cell labeling respective to background (see below).

B.4 Mapping

We analyzed sections between 1100 and 1300 μm from the midline; this particular level was chosen because it showed ZENK expression in response to all stimuli used. Human error in the identification of labeled cells was minimized by employing a computerized mapping system that applies a single and consistent set of criteria for labeling [Cecchi 98]. Briefly, sections were scanned using a computer-controlled microscope setup. High-resolution images were

acquired through a CCD camera (3 pixels/μm), and automatically pasted to reconstitute a composite image of each entire NCM section. For each pixel we measured the optical density (O.D.) resulting from DAB/nickel precipitation. Labeling intensity was defined as the ratio between the average O.D. of labeled nuclei and that of the background. Background was computed as the average O.D. of all tissue in NCM that was not recognized as labeled cells. The composite images were analyzed for recognition of labeled cell nuclei using criteria of size (3 to 10 μm diameter), eccentricity (ratio of elliptic radia $\leq$0.3), uniformity of labeling (variance of labeling over mean labeling $\leq$0.5), and shape (ratio between actual perimeter and perimeter of approximating ellipse $\leq$1.5). A lower cutoff for mean cell labeling intensity was set at 40 % above background tissue labeling, equivalent to two STD above average background. The rostral boundary of NCM was defined by drawing a straight line along the major axis of field L2a on parasagittal sections; the dorsal, ventral and caudal boundaries of NCM were naturally defined by the ventricular zone (Figure A.1B). In order to obtain an anatomical reference for comparison across maps, a standardized NCM outline was generated by aligning the outlines of all 80 sections relative to each other, and computing the average outline. All sections were then automatically aligned and scaled respective to this standardized NCM outline. The population of all cells studied (36,000 cells from 80 animals) was divided into three bins according to their labeling intensities, so that each bin contains a third of the total

cell population (low, 1.40 to 1.62; medium, 1.62 to 1.77; high, 1.77 to 2.68). To obtain average density maps of each stimulus group, the set of points of each map was convolved with a gaussian kernel of 50 μm radius, and an average map was generated for every 4 animals presented with the same stimulus. Information about labeling intensities (blue, green and red bins) was processed independently and finally superposed. Brightness in the maps as finally displayed represents absolute number of labeled cells per unit area, while hues display the relative proportions of cells with low, intermediate and high amounts of *ZENK* protein.

B.5 Clustering Analysis

To measure the inhomogeneity of the spatial distribution of cells in each pattern, we computed the index of clustering given by Thiel's redundancy measure (T) [Okabe 92]. For any pattern of scattered points (labeled cells in our case), T is defined as the negative of the difference between the Shannon entropy of the given pattern (S) and that of a pattern with equal number of points forming a regular grid (S_r), so that $T = -(S - S_r)$. For the calculation of Shannon entropy, each pattern was subjected to a Delauney triangulation [Okabe 92], and the entropy was computed as $S = -\sum p_i \log p_i / \log 2$, where p_i is the area of each triangle (A_i) relative to the total area (A), i.e. A_i/A. This measure is independent of the number of cells and does not take into account their labeling intensities (high/medium/low channels), thus being

sensitive only to the relative spatial distribution of cells [Okabe 92].

B.6 Sum Maps

To generate a simple sum map (Figure A.6B), we performed a pixelwise superposition of two whistle maps (b and d in Figure A.2), and subtracted from it the silence map (Figure A.1C). To generate the weighted sum map (Figure A.6C), we took the same two whistle maps and computed the linear combination that best fits the sequence map. The fit is a least square approximation of the map $f' = (\alpha b + \beta d - z)$ to the map f, where b and d are whistle maps and f the sequence map in Figure A.2, and z is the silence map in Figure A.1C. The coefficients found were $\alpha = 0.58$ and $\beta = 0.36$. To quantitate the differences between the map of a sequence and the sum maps, we used a similarity index that quantifies the proximity or similarity, pixelwise, of two given maps [Chapman 96]. It is computed as follows:

$$S = 1 - \sqrt{\frac{\sum_{i=0}^{N} (X_i - Y_i)^2}{N\Delta^2}}$$

where X_i and Y_i are the i^{th} pixel values, N is the total number of pixels, and Δ is the maximum value per pixel. This similarity index ranges from 0 to 1, corresponding to completely decorrelated maps and to identical maps, respectively. For comparison, the similarity index between two random maps (in which each i^{th} pixel value varies randomly) is 0.593.

B.7 PCA

To perform a Principal Component Analysis of the set of average maps, individual maps were parceled in space-bins, defined by a 6×6 grid. Each space-bin had three color bins, according to the labeling intensities of cells (see Mapping section above). This defined a 6×6×3 vector space, so that each map was represented by a 108 dimensional vector containing coarse-grained information on spatial distribution and labeling intensity of cells. The value of each vectorial component was computed as the number of cells falling within the corresponding space-color bin.

Bibliography

[Adjari 93] A. Ajdari and J. Prost, *C. R. Acad. Sci. Paris II*, **315** 1635 (1993).

[Adleman 94] Adleman, L. M. (1994). Molecular Computation of Solutions to Combinatorial Problems, *Science* **266**, 1021-1024.

[Adrian 28] Adrian, E. D. (1928). *The Basis of Sensation*, Norton, New York.

[Alberts 89] Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and J. D. Watson (1989). *The Molecular Biology of the Cell*, (2^{nd} edition), Garland, New York.

[Alvarez-Buylla 88] Alvarez-Buylla, A., and Vicario, D. S. (1988). Simple microcomputer system for mapping tissue sections with the light microscope, *J. Neurosci. Meth.* **25**, 165-173.

[Arieli 96] A. Arieli, A. Sterkin, A. Grinvald & A. Aertsen (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses, *Science*, **273** 868-871.

[Armstrong 65] Armstrong, C. M. and Binstok, L. *J. Gen. Physiol.* **48** 855-872 (1965).

[Armstrong 77] Armstrong, C. M. and Bezanilla, F. *J. Gen. Physiol.* **70** 577-590 (1977).

[Ash 90] R.B. Ash, *Information Theory*, Dover, New York (1990).

[Astumian 94] D. Astumian and M. Bier, *Phys. Rev. Lett.* **72** 1766 (1994).

[Athreya 72] Athreya, K.B. & Ney, P.E. (1972) *Branching Processes* (Springer-Verlag, Berlin).

[Atick 92] Atick, J.J. & Redlich, A.N. (1990). Towards a theory of early visual processing, *Neural Comp.* **4** 559-572.

[Ball 97] Ball, G.F., Tiemçiani, O., and Balthazart, J. (1997). Induction of the ZENL protein after sexual interactions in the male Japanese quail, *Neuroreport* **8**, 2965-2970.

[Barlow 61] Barlow, H.B. (1961). Possible principles underlying the transformation of sensory messages. In *Sensory Communication.* W. Rosenblith, ed. (MIT Press, Cambridge MA).

[Barlow 72] Barlow, H. B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology, *Perception*, **1** 371-394.

[Beaver 95] Beaver, D. (1995). Computing with DNA. *J. of Comp. Biol.* **2**, 1-7.

[Benkovic 87, 91] Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A. & Benkovic, S. J. (1987), *Biochemistry* **26** 8410-8417. Dahlberg, M. E. & Benkovic, J. (1991), *Biochemistry* **30**, 4835-4843. Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A. & Benkovic, S. J. (1987), *Biochemistry* **26** 8410-8417.

[Bennett 82] Bennett, C. H. (1982). The thermodynamics of computation, *Int. J. Theor. Phys.* **21**, 905-940.

[Bickerton 96] Bickerton, D. (1996). *Language and Human Behavior*, University of Washington Press, Seattle.

[Borg-Graham 98] Borg-Graham, L. J., Monier, C., Fregnac, Y. (1998). *Nature* **393** (6683), 369.

[Bourdieu 95] Bourdieu, L., Magnasco, M. O., Winkelman, D. A., and Libchaber, A. (1995). Actin filaments on myosin beds: The velocity distribution, *Phys. Rev. E* **52** (6) 6573-79.

[Brillouin 89] Brillouin, L. (1989). Life, theormodynamics and cybernetics. In *Maxwll's Demon: Entropy, Information and Computing*, H. Leff and A. Rex. eds. (Princeton, NJ: Princeton University Press).

[Brown 75] Brown, J.L. (1975). The evolution of auditory communication. In *The Evolution of Behavior* (New York: W.W. Norton).

[Bryant 76] Bryant, H. L. & Segundo, J. P. *J. Physiol. (London)* **260**, 279 (1976).

[Capsius 96] Capsius B., and Leppelsack H. J. (1996). Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird, *Hearing Res.* **96**, 59-70.

[Capson 92] Capson, T.L., Paliska, J.A., Fenn Kaboord, B., West Frey, M., Lively, C., Dahlberg M. & Benkovic, S.J. (1992) *Biochemistry* **31**, 10984-10994.

[Caramazza 96] Caramazza, A. (1996). Neuropsychology: The brain's dictionary, *Nature* **380**, 485-486.

[Carr 88] Carr, C. E., Konishi, M. (1998). Axonal delay lines for time measurement in the owl's brainstem, *Proc. Natl. Acad. Sci. USA* **85** 8311-8315.

[Cecchi 96] G.A. Cecchi & M.O. Magnasco (1996). Negative Resistance and Rectification in Brownian Transport, *Phys. Rev. Lett.* **76** (11), 1968-1971.

[Cecchi 98] Cecchi, G. A., Ribeiro, S., Mello, C. V., and Magnasco, M. O. (1999). An automated procedure for the mapping and quantitative anal-

ysis of immunocytochemistry of an inducible nuclear protein, in press, *J. Neurosci. Meth.*.

[Cecchi 99] G.A. Cecchi, M. Sigman, J-M. Alonso, L. Martínez, D.R. Chialvo & M.O. Magnasco (1999). Noise in neurons is message-dependent, submitted to *Phys. Rev. Lett.* (1999)

[Chaitin 75] Chaitin, G. J. (1975). A theory of program size formally identical to information theory, *J. Assoc. Comput. Mach.* **22**, 329-340.

[Chapman 96] Chapman, B., Stryker, M.P., Bonhoffer, T. (1996). Development of orientation preference maps in ferret primary visual cortex, *J. Neurosci.* **16**, 6443-6453.

[Chaudhuri 97] Chaudhuri, A. (1997). Neural activity mapping with inducible transcription factors, *NeuroReport* **8**, III.

[Chew 95] Chew, S.J., Mello C., Nottebohm F., Jarvis E., and Vicario DS. (1995). Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3406-3410.

[Chew 96a] Chew, S.J., Vicario, D. S., and Nottebohm, F. (1996a). A large-capacity memory system that recognizes the calls and songs of individual birds, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1950-1955.

[Chew 96b] Chew, S.J., Vicario, D. S., and Nottebohm, F. (1996b). Quantal duration of auditory memories, *Science* **274**, 1909-1914.

[Chomsky 57] Chomsky, N. (1957). *Syntactic structures*, Mouton Publishers, The Hague.

[Chomsky 59] Chomsky, N. (1959). On certain formal properties of grammar, *Information and Control* **2**, 137-167.

[Christy 88] Christy, B., Lau, L., and Nathans D. (1988). A gene activated in mouse 3T3 cells by serum factors encodes a protein with "zinc finger' sequences. *Proc. Natl. Acad. Sci. USA* **85**, 7857-7861.

[Christy 89] Christy, B., and Nathans D. (1989). DNA binding site of the growth factor-inducible protein Zif268, *Proc. Natl. Acad. Sci. USA* **86**, 8737-8741.

[Clementi 93] Clementi, M., Menso S., Bagnarelli, P., Manzin, A., Valenza, A. & Varaldo, P. (1993) *PCR Methods Applic.* **2**, 191-196.

[Cole 89] Cole, A.J., Saffen, A. J., D. W., Baraban J. M., and Worley P. F. (1989). Rapid increase of an immediate early gene messenger RNA in hippocampal neurons by synaptic NMDA receptor activation, *Nature* **340**, 474-476.

[Cormen 94] Cormen, T. H., Leiserson, C. E., and Rivest, R. L. (1994). *Introduction to Algorithms*, The MIT Press.

[de Ruyter 97] de Ruyter van Steveninck, R.R., Lewen G. D. Strong S. P. Koberle R. and Bialek W. (1997). Reproducibility and variability in neural spike trains, *Science* **275**, 1805-1808.

[Damasio 96] Damasio, H., Grabowski, T.J., Tranel, D., Hichwa, R.D., Damasio, A.R. (1996). A neural basis for lexical retrieval, *Nature* **380**, 499-505.

[De Koninck 97] P. De Knonick, H. Schulman (1997). Sensitivity of CaM Kinase II to the Frequency of Ca sup 2+ Oscillations, *Science*, **279** 227-230.

[Deadwyler 97] Deadwyler, S. A. and Hampson, R. E. (1997). The significance of neural ensemble codes during behavior and cognition. *Annu. Rev. Neurosci.* **20**, 217-244.

[Doering 94] C. Doering, W. Horsthemke and J. Riordan, *Phys. Rev. Lett.* **72**, 19, 2984 (1994).

[Dolmetsch 97] R.E. Dolmetsch, R.S. Lewis, C.C. Goodnow, J.I. Healy (1997). Differential activation of transcription factors by Ca sup 2+ response amplitude and duration, *Nature*, **386** 855-858.

[Dolmetsch 98] R.E. Dolmetsch, K. Xu & R.S. Lewis (1998). Calcium oscillations increase the efficiency and specificity of gene expression, *Nature*, **392** 933-936.

[Doupe 97] Doupe, A. (1997). Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development, *J. Neurosci.* **17**, 1147-1167.

[Dragunow 96] Dragunow, M. (1996). A role for immediate-early transcription factors in learning and memory, *Behav. Gen.* **26**, 293-299.

[Dubois 83] Dubois, J. M., Schneider, M. F. and Khodorov, B. I., *J. Gen. Physiol.* **81** 829-844 (1983).

[Durbin 90] R. Durbin, R. & G. Mitchison, G. (1990). A dimension reduction framework for understanding cortical maps, *Nature*, **343** 644-647.

[Edelsbrunner 83] Edelsbrunner, H., Kirkpatrick, D. G. and Seidel, R. (1983). On the shape of a set of points in the plane, *IEEE Transactions on Information Theory* **IT-29**(4), 551-559.

[Feller 68] Feller, W. (1968) *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd Ed. (Wiley, New York).

[Ferre 92] Ferre, F (1992) *PCR Methods Applic.* **2**, 1-9.

[Földiák 95] Földiák, P. & Young, M.P. (1995). In *The Handbook of Brain Theory and Neural Networks*, M.A. Arbib, ed., MIT Press.

[Georgopoulos 95] Georgopoulos, A.P. (1995). Reaching: coding in the motor cortex, in *The Handbook of Brain Theory and Neural Networks*, pages 783-787, M.A. Arbib, ed., MIT Press.

[Gestland 93] Gestland, R. & Atkins, J. (eds.), (1993) *The RNA world* (Cold Spring Harbor Press).

[Gibbs 90] Gibbs, R. A. (1990) *Analytical Chemistry* **62**, 1202-1214.

[Godard 91] Godard, R. (1991). Long-term memory of individual neighbours in a migratory songbird, *Nature* **350**, 228-229.

[Goelet 86] Goelet, P., Castelucci, V., Schacher, S., and Kandel, E. (1986). The long and the short of long-term memory T a molecular framework, *Nature* **332**, 419-422.

[Goldsmith 84] Goldsmith, A.R., Burke, S. and Prosser, J.M. (1984). Inverse changes in plasma prolactin and LH concentrations in female canaries after disruption end reinitiation of incubation, *J. Endocrinol.* **103**, 251-256.

[Goller 96] Goller, F. and Suthers, R.A. (1996). Role of syringeal muscles in controlling the phonology of bird song, *J. Neurophysiol.* **76**, 287-300.

[Gross 92] Gross, C.G., Representation of visual stimuli in inferior temporal cortex, *Phil. Trans. R. Soc. Lond. B.* **335**, 3-10 (1992).

[Guckenheimer 83] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, Berlin (1983).

[Guilliland 90] Guilliland, G., Perrin, S., Blanchard, K. & Bumm, F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2725-2729.

[Gupta 91] Gupta, M. P., Gupta, M., Zak, R., and Sukhatme, V. P. (1991). Egr-1, a serum-inducible zinc-finger protein, regulates transcription of the rat cardiac a-myosin heavy chain gene. *J. Biol. Cem.* **266**, 12813-12816.

[Güttinger 85] Güttinger, H.R. (1985). Consequences of domestication in the song structures in the canary, *Behavior* **94**, 254-278.

[Hagiwaram 76] Hagiwara, S., Miyazaki, S. and Rosenthal, N. P., *J. Gen. Physiol.* **67** 621-638 (1976).

[Harris 63] Harris, T. E. (1963) *The theory of branching processes* (Springer-Verlag, Berlin).

[Hartley 90] Hartley, R.S. (1990). Expiratory muscle activity during song production in the canary, *Resp. Physiol.* **81**, 177-187.

[Heil 91] Heil, P. and Scheich, H. (1991). Functional organization of the avian auditory cortex analogue. I. Topographic representation of isointensity bandwidth, *Brain Res.* **539**, 110-120.

[Hibbard 96] Hibbard L S., McCasland J. S., Brunstrom J. E., Pearlman A. L. (1996). Automated recognition and mapping of immunolabelled neurons in the developing brain, *J. Microscopy* **183**, 241-256.

[Hille 75a] Hille, B. (1975) J. Gen. Physiol. **66** 535-560.

[Hille 75b] Hille, B., Woodhull, A. M. and Shapiro, B. I. (1975) *Phil. Trans. R. Soc. Lond. B* **270** 301-318.

[Hille 78] Hille, B. and Schwartz, W., *J. Gen. Physiol.* **72** 409-442 (1978).

[Hille 84] Hille, B., *Ionic Channels of Excitable Membranes*, Sunderland MA: Sinauer Associates (1984).

[Hjelmfelt 92] A. Hjelmfelt & J. Ross (1992). Chemical implementation and thermodynamics of collective neural networks, sl Proc. Natl. Acad. Sci. USA **89**, 388-391.

[Hodgkin 52] Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol. (Lond.)* **116**: 449-472, 473-496; 497-506; **117**: 500-544.

[Honeycutt 92] Honeycutt, R. L. *Phys. Rev. A* **45**.2 600-603 (1992).

[Honkaniemi 95] Honkaniemi, J., Sagar S. M., Pyykonen I., Hicks K. J., and Sharp F. R. (1995). Focal brain injury induces multiple immediate early genes encoding zinc finger transcription factors, *Mol. Brain Res.* **28**, 157-163.

[Hopfield 95] J.J. Hopfield (1995). Pattern recognition computation using action potential timing for stimulus representation, *Nature* **376** 33-36.

[Hose 87] Hose, B., Langner, G., Scheich, H. (1987). Topographic representation of periodicities in the forebrain of the mynah bird: one mapo for pitch and rythm?, *Brain Res.* **422**, 367-373.

[Immelman 69] Immelman, K. (1969). Song development in the Zebra Finch brain and other estrildid finches. In *Birds Vocalizations, Hinde*, R.A. Ed. (Cambridge Univ. Press, London, UK), 61-74.

[Invitrogen] Kit from Invitrogen (San Diego, CA).

[Iris Prog.] Iris Media Library Programming Guide, Silicon Graphics Inc. (1994).

[Iyer 99] Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C. F., Trent, J. M., Staudt, L. M., Hudson Jr., J., Boguski, M. S., Lashkari, D., Shalon, D., Botstein, D., and Brown, P. O. (1999). The Transcriptional Program in the Response of Human Fibroblasts to Serum, *Science* **283**, 83-87.

[Jarvis 97a] Jarvis E. D., Schwabl H., Ribeiro S., and Mello C. V. (1997). Brain gene regulation by territorial singing behavior in freely ranging songbirds. *Neuroreport* **8**, 2073-2077.

[Jarvis 97b] Jarvis E. D., Nottebohm F. (1997). Motor-driven gene expression, Proc. Natl. Acad. Sci. **94**, 4097-4102.

[Jin 97] Jin H., Clayton D.F. (1997). Localized changes in immediate early gene regulation during sensory and motor learning in zebra finches, *Neuron* **19**, 1049-1059.

[Joerges 97] Joerges, J., Kuttner, A., Galizia, C. G., and Menzel, R. (1997). Representations of odours and odour mixtures visualized in the honeybee brain, *Nature* **387**, 285-288.

[Joyce 94] G. F. Joyce (1994). In-Vitro Evolution of Nuclei-Acids, *Curr. Opin. Struct. Biol.* **4** 331-336.

[Kahn 97] D. Kahn, E.F. Pace-Schott, J.A. Hobson (1997). Consciousness in waking and dreaming: the roles of neuronal oscillation and neuromodulation in determining similarities and differences, *Neurosci.* **78** 13-38.

[Kaplan 96] P.D. Kaplan, G.A. Cecchi, A.J. Libchaber, unpublished.

[Kaplan 98] P. Kaplan, G.A. Cecchi and A.J. Libchaber (1998). DNA Based Molecular Computation: Template-Template Interactions in PCR, *DNA Based Computers II*, DIMACS: Series in Discrete Mathematics and Theoretical Computer Science, Vol **44**, American Mathematical Society (1998).

[Khachigian 96] Khachigian, L.M., Lindner, V., Williams A. J., and Collins T. (1996). Egr-1-induced endothelial gene expression: a common theme in vascular injury, *Science* **271**, 1427-1431.

[Knight 72] B.W. Knight, *J. Gen. Physiol.* **59**, 734 (1972).

[Koch 98] C. Koch, *Biophysics of computation*, Oxford University Press, New York (1998).

[Kohonen 96] T. Kohonen (1996). *Self Organizing Maps*, Springer, New York.

[Kolmogorov 65] Kolmogoroff, A. N. (1965). Three approaches to the quantitative definition of information, *Problmes Inform. Transmission* **1**, 1-7.

[Konishi 91] Konishi, M. (1991). Deciphering the Brain's Codes, *Neural Computation* **3**, 1-18.

[Kornacke 69] Kornacker, K., in *Biological Membranes*, Dowben, R. M, ed., Little, Brown and Co., Boston (1969).

[Kroodsma 96] Kroodsma, D.E. and Miller, E.H. (1996). *Ecology and Evolution of Acoustic Communication in Birds* (Cornell Univ. Press, New York).

[Kugler 90] Kugler P. (1990). Enzyme histochemical methods applied in the brain, *Eur. J. Morphol.* **v. 28**, 109-120.

[Lashley 50] Lashley, K.S. (1950). In search of the engram, *Symp. Soc. Exp. Biol.* **4** (1950).

[Laurent 96] M. Wehr & G. Laurent (1996). Odour encoding by temporal sequences of firing in oscillating neural assemblies, *Nature* **384** 162-166.

[Leff 94] Leff, H.S. & Rex. A.F., *Maxwell's demon: information, entropy, computing*, A. Hilger (Europe) and Princeton U.P. (USA) (1994).

[Lemaire 88] Lemaire, P., Revelant, O., Bravo, R., and Charnay, P., Two mouse genes encoding potential transcription factors with identical DNA-binding domains are activated by growth factors in cultured cells, *Proc. Natl. Acad. Sci. USA* **85**, 4691-4695 (1988).

[Leppelsack 76] Leppelsack, H.J. and Vogt, M. (1976). Responses of auditory neurons in the forebrain of a songbird to stimulation with species-specific sounds, *J. Comp. Physiol. [A]* **107**, 263-274.

[Levin 74] Levin, L. A. (1974). Laws of information conservation and aspects of the foundation of probability theory, *Problems Inform. Transmission* **10**, 206-210.

[Li 93] Li, M., Vitányi, P. (1993). *An introduction to Kolmogorov complexity and its applications*, Springer-Verlag, New York NY.

[Lilley 95] D. M. Lilley (1995). Kinking of DNA and RNA by base bulges, *Proc. Natl. Acad. Sci. USA* **92** 7140-7142.

[Lipton 95] R. J. Lipton (1995). DNA Solution of Hard Computational Problems, *Science* **26** 542-545.

[Longtin 98] A. Longtin & D.R. Chialvo, *Phys. Rev. Lett.* **81**, 4012 (1998).

[Lopatin 94] Lopatin, A. N., Makhina, E. N. and Nichols, C. G., *Nature* **327** 366-369 (1994).

[Müller 85a] Müller, C.M. and Leppelsack, H.J. (1985). Feature extraction and tonotopic organization in the avian auditory forebrain, *Exp. Brain. Res.* **59**, 587-599.

[Müller 85b] Müller, C.M. and Scheich, H. (1985). Functional organization of the avian auditory field L: a comparative 2DG study, *J. Comp. Physiol.* **156**, 1-12.

[Müller 88] Müller, C.M. (1988). Distribution of GABAergic perikaria and terminals in the centers of the higher auditory pathway of the chicken, *Cell Tissue Res.* **59**, 587-599.

[Magnasco 93] M. Magnasco, *Phys. Rev. Lett.* **71**.10, 1477 (1993).

[Magnasco 94] M. Magnasco, *Phys. Rev. Lett.* **72**.16, 2656 (1994).

[Magnasco 97] M. Magnasco (1997). Chemical Kinetics is Turing Universal, sl Phys. Rev. Lett. **78**, 1190-1193.

[Mainen 95] Z.F.Mainen & T.J. Sejnowski, *Science* **268** (5216), 1503 (1995).

[Margoliash 92] Margoliash, D. and Fortune, E. S. (1992). Temporal and harmonic combination-sensitive neurons in the zebra finch's HVc, *J. Neurosc.* **12**, 4309-4326.

[Margoliash 97] Margoliash D. (1997). Functional organization of forebrain pathways for song production and perception, *J. Neurobiol.* **33**, 671-93.

[Marler 82] Marler, P. and Peters, S. (1982). Sparrows learn adult song and more from memory, *Science* **213**, 780-782.

[Martínez 98] Martínez, L. M. & Alonso, J. M. (1998). *Nature Neuroscience*, **1**,5:395-403.

[McAdams 95] McAdams, H. H. & Shapiro, L. (1995). Circuit simulation of genetic networks, *Science* **269**, 650-656.

[McCulloch 43] McCulloch, W.S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity, *Bulletin of mathematical biophysics*, **5** 115-133.

[Mello 92] Mello, C.V., Vicario D.S., and Clayton D. F. (1992). Song presentation induces gene expression in the songbird forebrain, Proc. Natl. Acad. Sci. U.S.A. **89**, 6818-6822.

[Mello 94] Mello C. V. & Clayton D. F. (1994). Song-induced ZENK gene expression in auditory pathways of songbird and its relation to the song control system, J. Neurosc. **14**, 6652-66.

[Mello 95] Mello, C.V., Nottebohm, F. Clayton, D. F. (1995). Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene's response to that song in zebra finch telencephalon, J. Neurosci. **15**, 6919-6925.

[Mello 98a] Mello, C.V. and Ribeiro, S. (1998). ZENK protein regulation by song in the brain of songbirds, *J. Comp. Neurol.* **393**, 426-438.

[Mello 98b] Mello, C. V., Vates, G. E., Okuhata, S., and Nottebohm, F. (1998). Descending auditory pathways in the adult male zebra finch (Taeniopygia guttata), *J. Comp. Neurol. 395*, 137-160.

[Milbrandt 87] Milbrandt, J. (1987). A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor, *Science* **238**, 797-799.

[Millonas 94] M. Millonas and M. Dykman, *Phys. Lett.* **A185** 65 (1994).

[Minorsky 62] N. Minorsky, *Nonlinear Oscillations*, Van Nostrand, Princeton (1962).

[Minsky 67] Minsky, M., *Computation: finite and infinite machines*, NJ: Prentice-Hall (1967).

[Moll 88] Moll, R., Arbib, M., Kfoury, A.J., *An introduction to formal language theory*, Springer-Verlag (1988).

[Morgan 89] Morgan, J. I., and Curran, T. (1989). Stimulus-transcription coupling of neurons: role of cellular immediate-early genes, *TINS* **12**, 459-462.

[Morgan 91] Morgan, D. P., Scofield, C. L. (1991). *Neural networks and speech processing*, Kluwer Academic Press, Norwel MA.

[Mountcastle 74] Mountcastle V.B. (Ed) *Medical Physiology*, C.V. Mosby (1974)

[Mullis 87] Mullis, K. B. & Faloona, F. A. (1987) *Methods Enzymol.* **155**, 335-350.

[Mundiger 95] Mundiger, P.C. (1995). Behavior-genetic analysis of canary song: inter-strain differences in sensory learning, and epigenetic rules, *Anim. Behav.* **50**, 1491-1511.

[Nastiuk 94] Nastiuk K. L., Mello C. V., George J. M., and Clayton D. F. (1994). Immediate-early gene responses in the avian song control system: cloning and expression analysis of the canary c-jun cDNA, *Brain Research. Molecular Brain Research* **27**, 299-309.

[Nicolelis 95] Nicolelis, M.A.L., Baccala, L. A., Lin R. C., and Chapin J. K. (1995). Sensorimotor encoding by synchronous neural ensemble activity at multiple levels of the somatosensory system, *Science* **268**, 1353-1358.

[Nicolelis 98] Nicolelis, M. A. L., Ghazanfar, A. A., Stambaugh, C. R., Oliveira, L. M. O., Laubach, M., Chapin, J. K., Nelson, T. J., & Kaas, J. H. (1998). Simultaneous encoding of tactile information by three primate cortical areas, *Nature Neurosci.*, **1** 621-630.

[Nishimura 88] Nishimura, T., and Vogt, P. (1988). The avian homologe of the oncogene jun, *Oncogene* **3**, 659-663.

[Nottebohm 72] Nottebohm, F. (1972). The origins of vocal learning, *The Amer. Nat.* **106**, 116-140.

[Nottebohm 78] Nottebohm, F. and Nottebohm, M.E. (1978). Relationship between song repertoire and age in the canary brain, Serinus canarius, *Z. Tierpsychol.* **46**, 298-305.

[Okabe 92] Okabe, A., Boots, A., and Sugihara, K. (1992). In *Spatial tessellation concepts and applications of Voronoi diagrams* (Wiley, New York).

[Olshausen 96] Olshausen, B.A. & Field, D.J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images, *Nature* **381** 607-9.

[Ouyang 97] Q. Ouyang, P. Kaplan, S. Liu & A. Libchaber (1997). DNA solution of the Maximal Clique Problem, *Science* **278**, 446-449.

[Patel 91] Patel, S.S., Wong, I. & Johnson, K.A. (1991) *Biochemistry* **30**, 511-525.

[Paun 98] Paun, G., Rozenberg, G., Salomaa, A., Brauer, W., eds. (1998).*DNA Computing : New Computing Paradigms*, Springer-Verlag, New York.

[Press 94] W. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery (1994). *Numerical recipes in C*, Cambridge University Press.

[Raeymakaers 93] Raeymaekers, L. (1993) *Annal Biochem.* **214**, 582-585.

[Rauschecker 95a] Rauschecker, J.P. (1995). Compensatory plasticity and sensory substitution in the cerebral cortex, *TINS*, **18**, 36-43.

[Rauschecker 95b] Rauschecker, J.P., Tian, B., Hauser, M. (1995). Processing of Complex Sounds in the Macaque Nonprimary Auditory Cortex, *Science* **268** 111-114.

[Reale 80] Reale, R. A. and Imig, T. J. (1980). Tonotopic organization in auditory cortex of the cat, *J. Comp. Neurol.*, **192**, 265-291.

[Reif 95] J. Reif (1995) Parallel Molecular Computation: Models and Simulation, *Proc. of 7th Annual ACM Symposium on Parallel Algorithms and Architectures SPAA '95*, 213-223.

[Ribeiro 98] Ribeiro, S., Cecchi, G. A., Magnasco, M. O., and Mello, C. V. (1998). Toward a song code: syllabic representation in the canary brain, *Neuron* **21**, 359-371 .

[Ribeiro 98] Ribeiro, S., Cecchi, G. A., Magnasco, M. O., and Mello, C. V. (1998). Ontogenesis of Sillabic Maps in NCM, *Soc. Neurosci. Abstr.* (1998).

[Richmond 87] Richmond, B.J. and Optican, L.M. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform, *J. Neurophysiol.* **57**, 147-161.

[Rieke 97] F. Rieke, D. Warland, R. de Ruyter van Steveninck, W. Bialek, *Spikes: exploring the neural code*, MIT Pres, Cambridge (1997).

[Risken 89] Risken, H. (1989) *The Fokker-Planck Equation: Methods of Solution and Applications*, Berlin: Springer-Verlag.

[Roelfsema 98] P.R. Roelfsema PR & W. Singer (1998). Detecting connectedness, *Cerebral Cortex* **8** 385-96.

[Rogers 67] Rogers, H., *Theory of recursive functions and effective computability*, McGraw Hill (1967).

[Rolls et al. 1997] Rolls, E.T., Treves A., and Tovee M. J. (1997). The representational capacity of the distributed encoding of information provided

by populations of neurons in primate temporal visual cortex, *Exp. Brain. Res.* **114**, 149-162.

[Salomaa 73] Salomaa, A. (1985). *Computation and automata*, Encyclopedia of Mathematics vol. 25, Cambridge University Press, Cambridge MA.

[Sambrook 89] Sammbrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning, a laboratory manual*, 2nd Ed.

[Scheich 91] Scheich, H. (1991). Auditory cortex: comparative aspects of maps and plasticity, *Curr. Opin. Neurobiol.* 1, 236-247.

[Seyfert 89] Seyfert, V.L., Sukhatme V. P., and Monroe J. G. (1989). Differential expression of a zinc finger-encoding gene in response to positive versus negative signaling through receptor immunoglobulin in murine B lymphocytes, *Mol. Cell. Biol.* **9**, 2083-2088.

[Shannon 48] Shannon, C.E. (1948). A Mathematical Theory of Communication, *Bell Sys. Tech. J.* **27**.

[Shannon 49] Shannon, C. and Weaver, W. (1949). *The mathematical theory of communication*, University of Illinois Press, Urbana.

[Solis 97] Solis, M. M. and Doupe, A. J. (1997). Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning, *J. Neurosci.* **17**, 6447-6462.

[Solomonoff 64] Solomonoff, R. J. (1964). A formal theory of inductive inference, part 1 and 2, *Inform. Contr.* **7**, 224-254.

[Stolovitzky 96] G. Stolovitzky & G.A. Cecchi (1996). Efficiency of DNA Replication in the Polymerase Chain Reaction, *Proc. Natl. Acad. Sci. USA* **93** (23), 12947-12952.

[Stripling 97] Stripling R., Volman S. F., and Clayton D. F. (1997). Response modulation in the zebra finch neostriatum: relationship to nuclear gene regulation, *J. Neurosc.* **17**, 3883-3893.

[Sukhatme 88] Sukhatme, V.P., Cao, X., Chang, L.C., Tsai-Morris, C.H., Stamenkovitch, D., Ferreira, P.C.P., Cohen, D.R., Edwards, S.A., Shows, T.B., Curran, T., LeBeau, M.M., and E.D. Adamson (1988). A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization, *Cell* **53**, 37-43.

[Sun 95] Sun, F. (1995) *J. Comp. Biol.* **2**, 63-86.

[Szostak 93] J. W. Szostak & A. D. Ellington (1993). In Vitro Selection of Functional RNA sequences. In R. F. Gesteland, J. F. Atkins, eds., *The RNA World*, Cold Spring Harbor Press, Cold Spring Harbor.

[Theurich 84] Theurich, M. Müller, C.M. and Scheich, H. (1984). 2-Deoxyglucose accumulation parallels extracellularly recorded spike activity in the avian auditory neostriatum, *Brain Res.* **322**, 157-161.

[Triggs-Raine 90] B. L. Triggs-Raine & R. A. Gravel (1990). Diagnostic heteroduplex: simple detection of carriers of a 4-bp insertion mutation in Tay-Sachs disease, *Am. J. Hum. Genet.*, 46:183-184.

[Turing 36] Turing, A.M. (1936). On computable numbers, with an application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society* **2(42)** 230-265.

[van Kampen 81] van Kampen, N. G. (1981) *Stochastic processes in physics and chemistry* (North-Holland, Amsterdam).

[von Neumann 56, 66] J. von Neumann, *Theory of self-reproducing automata*, A. W. Burke, ed. Univ. of Illinois Press (Urbana, IL 1966). *Probabilistic logic and the synthesis of reliable organisms from unreliable components*, C. E. Shannon and J. McCarty, ed. Princeton Univ. Press (Princeton, NJ 1956).

[Vallet 95] Vallet, E.M. and Kreutzer, M.L. (1995). Female canaries are sexually responsive to special song phrases, *Anim. Behav.* **49**, 1603-1610.

[Vallet 92] Vallet, E.M., Kreutzer, M.L., and Richard, J-P. (1992). Syllable phonology and song segmentation: testing their salience in female canaries, *Behavior* **121**, 155-167.

[Vates 96] Vates, G.E., Broome B. M., Mello C. V., and Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation

to the song system of adult male zebra finches, *J. Comp. Neurol.* **366**, 613-642.

[Vicario 93] Vicario, D.S. & Yohay, K.H. (1993).*J. Neurobiol.*, **24** 488-505.

[Wallace 95] Wallace, C. S., Withers G. S., Weiler I. J., George J. M., Clayton D. F., and Greenough, W. T. (1995). Correspondence between sites of NGFI-A induction and sites of morphological plasticity following exposure to environmental complexity, *Mol. Brain Res.* **32**, 211-220.

[Weiss 95] Weiss, G. & von Haeseler, A. (1995) *J. Comp. Biol.* **2**, 49-61.

[Wingfield 90] Wingfield, J.C. and Goldsmith, A.R. (1990). Plasma levels of prolactin and gonadal steroids in relation to multiple-brooding and renesting in gree-living populations of the song sparrow, Melospiza melodia, *Horm. Behav.* **24**, 89-103.

[Woolsey 42] Woolsey, C.N. and Walzl, E.M. (1942). Topical projection of nerve fibers from local regions of the cochlea to the cerebral cortex of the cat, *Bull. Johns Hopkins Hosp.* **71**, 315-344.

[Yin 95] Yin, H., Wang, M. D., Svoboda, K., Landick, R., Block, S. M. & Gelles, J. (1995) *Science* **270**, 1653-1657.

[Zurek 88] Algorithmic information content, Church-Turing Thesis, physical entropy, and Maxwell's Demon, in *Complexity, entropy and the physics*

of computation, Santa Fe Institute studies in the science of complexity vol. III, Addison-Wesley, Redwood City CA.

End