STUDIES OF THE IMMUNE NETWORK BASED ON SHAPE-SPACE AND DISTANCE COEFFICIENT

By

Sophie Royer

B. Sc. (Physique) Université de Montreal, 1991

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTERS OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF PHYSICS

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December 1993

© Sophie Royer, 1993

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Physics

The University of British Columbia
Vancouver, Canada

Date December 9th, 1993

DE-6 (2/88)

Abstract

A selective history of immunology is first presented that tells about the development of the theories of "acquired immunity", from the early observations of the phenomenon of immunity in the Roman times until the long debate ending in the late 1960's between the proponents of the theories of "cellular immunity" and those of the theories of "humoral immunity". Some theories of antibody formation are also reviewed up to the "clonal selection theory". Then the "immune network hypothesis" and some of the models that it engendered are explained, with a focus on a particular model due to Hoffmann and his co-workers: the N-dimensional network model. A short history of the attempts to model the "affinity" distribution and various choices of "connectivity" matrices is presented, focussed on a particular one: a connectivity matrix based on a one-dimensional "shape-space". The $\pm$ shape-space [1], a shape-space formulated by Segel and Perelson, is reviewed and a "new shape-space without shape zero", the Δ shape-space, is introduced in which "complementarity" relationships between clones differ from the ones in the $\pm$ shape-space. Some analysis and numerical simulations of the two different versions of shape-space embedded in Hoffmann's N-dimensional network are shown, which are the first simulations of the model to have ever been be done with non-Boolean affinities. The concept of a "distance coefficient" is reviewed, analyzed and developed further for its use with non-Boolean affinities. The first numerical simulations of the distance coefficients to have ever been done are presented and analyzed, embedded in Hoffmann's N-dimensional network model with Boolean and non-Boolean affinities.

[1] The author's renaming of the original shape-space of Segel and Perelson.

Table of Contents

List of Figures

Acknowledgements

There are so many people that I could thank for their support and friendship throughout my stay in Vancouver and the completion of this thesis, but more particularly I am very grateful to those that follow here.

I would like to thank my supervisor Geoffrey Hoffmann for having had so much patience during the first months while I was beginning to familiarize myself with computers and trying to learn what I could about theoretical biology, and for his continuing support throughout the entire degree. I also have to thank Leah Keshet who always had faith in me.

I would like to thank my father Antoine Royer and mother Bach-Tuyet Vo for always having had so much faith in me and for having helped me in many ways. I also hope that my grand-father Jacques Royer who deceased toward the end of the completion of this thesis is proud of me, wherever he is now. He was also a great source of support and of inspiration.

My very dear friends that have helped me in one way or another and that I should thank more particularly are Robin Barley, Gul Civelekoglu and Cord Seele.

I must also thank those that have shared their knowledge on computer science with me: Alan Boulton, Roger Kemp and Henry Lee.

Chapter 1

Introduction

Immunology is among the most dynamical branches of biology. An amazing number of facts are being discovered every year. But still it is realized that a lot is left to understand about the immune system, and especially about its regulation.

The history of immunology starts a long time ago, with observations of the phenomenon of "acquired immunity", that is observations that those that have recovered from a disease are "immune" to it upon a second attack. Chapter 1 of this thesis begins with a history of those early observations and presents the first theories of acquired immunity up to the long battle between supporters of "cellular immunity" who thought that cells were the important constituents responsible for the defense of the immune system, and the supporters of "humoral immunity" who considered "antibody" molecules to be the defenders of the body. It also presents the development of the theories of antibody formation up to the "theory of clonal selection", which is now considered to be a cornerstone in immunology. The information contained in Chapter 1 is mostly derived from the book titled "A history of immunology" [125] by Arthur Silverstein.

The first mathematical models in immunology that would try to give insight to the question of the regulation of the immune system were essentially triggered by the "network theory" of self-regulation. Chapter 2 presents the network hypothesis and reviews the first models of immune networks, focussing on a particular one: Hoffmann's model. It also mentions some later work by other theorists in the field. It then presents Hoffmann's "N-dimensional network" model which is the basis of the work presented in Chapters 4

and 5.

Chapter 3 talks about a practical aspect that network theorists face when they want to input the biological parameters "affinities" in their models. It tells about attempts to model the "affinity distribution" of antibodies and the eventual choices of "connectivity matrices" (matrices of affinities) that network theorists have made.

Chapter 4 presents some of the results obtained when implementing Hoffmann's N-dimensional network into a particular setting of computed affinities presented in chapter 3, the "shape-space". It first shows a few simulations done with the original shape-space of Segel and Perelson which was the earliest version of shape-space. Then it presents a new shape-space without shape-zero and shows some simulations done with it.

Chapter 5 contains the most analytical and principal part of the thesis. It presents a generalization of the "distance coefficient" which is a measure of the "dissimilarity" between two substances for its use with non-Boolean affinities. Some simulations are also shown which are based on Hoffmann's N-dimensional network model with affinities of the new shape-space without shape zero. It can actually be said that all previous chapters to this one are only introductory to it. In order to justify the generalization of the distance coefficient for its use with non-Boolean affinities and to show some simulations of it, there was a need to use an immune network model, so the N-dimensional network model needed to be presented. It then seemed a good idea to present the network hypothesis which originated the model, as well as the background to the hypothesis. Also, since the goal of the generalization was to enable the use of non-Boolean affinities in the concept of a distance coefficient. it seemed sensible to mention the problem of the choice of the affinity distribution. The affinity matrix chosen here being of a shape-space, it also made sense to show simulations in shape-space with Hoffmann's model, since the model had always been used with Boolean affinities.

Finally, chapter 6 contains the conclusion.

The simulations that are presented in this thesis have been done with Fortran programs created by the author, excepted for some Fortran routines obtained from the book "Numerical Recipes in Fortran" [113]. For the integration of the differential equation of Hoffmann's N-dimensional network model, the following routines from chapter 16 were used: the stepper routine "rkqs"which is a fifth-order Runge-Kutta step with monitoring of the local truncation error to ensure the accuracy and adjust the stepsize, the algorithm routine "rkck" which is a fifth-order Cash-Karp Runge-Kutta method, the driver routine "odeint" with adaptive stepsize control (modified by the author). For the generation of random initial conditions (variables having a value between 0 and 1), the routine "Ran2" from chapter 7 was used. It is a long period ($> 2 \times 10^{18}$) random number generator of l'Ecuyer with Bays-Durham shuffle and added safeguards.

The original illustrations that are part of this work were made with the graphics packages: Gnuplot, MG, and Mathematica.

Chapter 2

From immunity to clonal selection

2.1 The phenomenon of immunity

The latin word "immunitas" is related to the legal concept of exemption [125]. It initially described the exemption of a person from service or duty in Rome. Later, in the Middle Ages, it referred to the exemption of the Church and its properties and personnel from civil control. The first use of the term in a disease context can be traced to the Roman poet Marcus Annaeus Lucanus (A.D. 39-65) when he observed the great resistance of the Psylli tribe of North Africa to snakebites. It was later used occasionally in the same manner, but gained popularity only in the nineteenth century, with the discovery of smallpox vaccination by Edward Jenner.

The phenomenon of acquired immunity had been observed long ago. Pestilence and poison were once the most feared causes of death. Even then, some keen observers realized that a person who had been infected by a disease and recovered would not be affected by it again, at least not fatally. The first related report comes from the historian Thucydides [139] of Athens, in 430 B.C. Also, it is known that in Roman times, Mithridates VI, king of Pontus, wrote in his medical commentaries about his increasing daily intakes of poison, to prevent himself from attempts against his life. His conqueror Pompey had those writings translated, and the practice became well-used, so that throughout the Middle Ages there were mixtures called Mithridaticum used to that effect. It is worthy to note that for centuries, it was thought that many diseases were due to poison, called

"virus" in latin. Also, the Greek word "pharmakeia" still means poisoning, witchcraft or medicine!

The first theory of acquired immunity was made by the best known of the Islamic physicians, Abu Bekr Mohammed ibn Zakariya al-Razi (880-932). In his *Treatise on the Smallpox and Measles* [115], he gave the first modern clinical description of smallpox, stated the fact that survival from smallpox conferred lasting "immunity" and provided an explanation for it [126].

The notions of a disease being caused by small seeds (seminaria) and of its contagion was first raised by Fracastoro [46] in 1546. It is interesting to note that he thought that those seminaria have affinities for different objects. And with this he explained "natural immunity" to certain diseases. Furthermore, he stated that the seminaria of smallpox have an affinity for that trace of menstrual blood that each individual was supposed to have inherited from its mother. Shortly after, Girolamo Mercuriali [92] denied the theory of menstrual blood affinity stating that if smallpox, measles, and leprosy were all results of menstrual blood contamination, then getting one disease should give protection from the others, which obviously was not the case. One can see there a resemblance to the modern concept of "cross-immunity".

By the end of the seventeenth century, smallpox (previously a rare disease) had become a serious infection, which was going to plague the world for long after. With the rise of physical sciences in the sixteenth and seventeenth centuries, new theories of smallpox were developed, the most dominant ones being those proposed by the iatrophysicists and the iatrochemists who would explain everything in terms of physical and chemical processes respectively [127].

At the beginning of the eighteenth century, the practice of variolation (consisting of inoculation of smallpox to produce mild dermal infections which prevent further more severe attacks), that had become popular as part of the folk medicine of several Eastern

cultures, gained particular attention from the official Western medicine, through the independent influences of Cotton Mather [88], a Boston resident, and of Lady Mary Montagu [49], the wife of a British Ambassador in Constantinople. Several interesting theories were raised as explanations for the phenomenon of "acquired immunity" against smallpox [128].

The humoral theory of diseases was then still prevalent, as a remnant of the last twenty centuries belief that diseases originated from a maladjustment of the four humors: the blood, the phlegm, the yellow bile and the black bile. It was only in 1858 that Rudolf Virchow [143] suggested that all pathology is based on the malfunction of cells rather than the maladjustment of humors. His successful theory gave birth to cellular pathology.

In the late 1870's, modern bacteriology was founded, mainly as a result of the work of Robert Koch [78] and Louis Pasteur [104]. They showed that each infectious disease is produced by its own specific pathogenic microorganism. Studies of these organisms both *in vitro* and *in vivo*, led to the germ theory of disease. Pasteur [103] was able to induce acquired immunity to chicken cholera, but without being able to suggest an adequate mechanism that would account for it.

Whether inflammatory responses were beneficial or detrimental to one's health was still controversial. Because inflammatory responses often seemed to accompany infectious diseases, the latter theory was most favored. As microscopy and anatomical pathology developed, macrophages and microphages were identified as the important cells present in inflammatory responses and were thought by most pathologists to carry the infectious organisms throughout the body. Then, in 1884, the Russian zoologist Metchnikoff [93] proposed that phagocytic cells constitute the primary line of defense of an organism.

He was immediately opposed by the supporters of humoral immunity, which were not able then to rigorously disprove him. But the discovery in 1888 by Nuttall [98] that the serum of normal antibodies has a natural toxicity for some microorganisms was the start

of further investigations in that direction. During the same year, Metchnikoff became *chef de service* at the Pasteur Institute where hostility toward scientists based at Koch's Institute in Berlin reigned. This feeling was inherited from a violent debate that took place between Louis Pasteur and Robert Koch since the late 1870's, caused by the international politics that had not long ago opposed France to Germany. Each accused the other of scientific incompetence and refuted the validity of the other's experiments. Thus, following this, a fierce and nationalist battle lasting for two decades ensued. Some scientists grouped around Metchnikoff and his phagocytic theory at the Pasteur Institute were in favour of cellular immunity. They claimed that the phagocytic and digestive powers of the macrophage and of the microphage constituted the main defense of the body against infection. Others based at the Koch's Institute in Berlin under the leadership of Koch were for humoral immunity. They argued that invading pathogens could be immobilized and destroyed only by the soluble substances of the blood and other body fluids. Each group performed experiments designed to support their own theory. Metchnikoff and his co-workers did numerous experiments to prove that the natural bactericidal powers of the sera of different species are often not related to the species' susceptibility to infection by a given organism. His opponents would continually report that bacteria could be killed by the cell-free fluids of normal and especially of immunized animals. The discovery in 1890 by von Behring and Kitasato [144] that immunity to diptheria and tetanus is due to antibodies without obvious interactions of any cellular elements hit hard the cellular theory of immunity. Von Behring and Wernicke [145] observed soon after that passive transfer of immune serum protected a healthy organism from diptheria, without any involvement of any cellular activity. Pfeiffer [111] then found direct bacteriolysis of cholera microorganisms by antibodies in guinea pigs, and shortly after it was found that this phenomenon could also be induced in a normal animal by injection of an immune serum. And Bordet [12] would soon observe that even the erythrocyte could be killed by

antibodies in the absence of phagocytes.

More observations were made over the 1890's that showed the importance of antibodies for immunity. As more antibodies of different specificities were being discovered and as Ehrlich [36] was able to isolate antibodies in a test tube via the recently discovered precipitin reaction [80], more immunologists turned towards the humoral theory, as they felt more comfortable with the antibodies than with the less manageable phagocytes. In 1908, the Swedish Academy tried to reconcile both camps by confering a joint Nobel prize to Metchnikoff and Ehrlich, then leader of the humoral theory "in recognition of their work in immunity". It is interesting to note that in England, Wright and Douglas [153] claimed that "both humoral and cellular functions were equally important and interdependent, in that humoral antibody interacts with its target microorganism to render it more susceptible to phagocytosis by macrophages" But apart from them, the theory failed to find strong supporters, and in the following decades only rarely would someone explore the role of cells in immunity or the phenomena of "bacterial allergy" and of "delayed sensitivity" which had been observed. In the 1920's and early 1930's, Zinsser [155, 156], investigated bacterial allergy and Dienes and co-workers [35] got interested in delayed hypersensitivity followed by two other groups [71, 130]. The importance of inflammatory cells in tuberculin allergy was shown by Rich and Lewis [116]. In the mid 1940's, Harris, Ehrich and co-workers [52] studied extensively the role of lymphocytes in antibody formation.

Only in the 1960's was there a switch from a chemical to a more biological approach to immunology. It was then realized that the phenomena of allograft rejection, immunological tolerance, immunity in some viral infections, pathogenesis of autoallergic diseases and those related to immunological deficiency diseases could not be explained within the framework of a theory based uniquely on humoral antibodies. This was the start of an explosion of research in cellular immunity that would try to catch up with the last

decades of inaction in that domain.

2.2 Theories of antibody formation

The discovery of humoral antitoxic antibodies in the early 1890's was the start of a theoretical investigation that would last for about eighty years. At that time, nothing was known about the nature of toxins or antitoxins and little was known about the chemistry of biological macromolecules in general. The questions that immunologists would try to answer were about the origin and formation of antibodies within the immunized host and about the way that they would acquire their specificity.

2.2.1 Antigen incorporation theories

The antigen incorporation theories appeared initially, with the first one proposed in 1893 by the noted German bacteriologist Hans Buchner [19]. He suggested that the antitoxin was derived from the antigen. But new experiments [121, 77] would soon disprove the hypothesis by showing that the amount of antibody formed in the immunized animal was far greater than the amount of antigen utilized. Despite this, there were several immunologists that continued to formulate theories to support the hypothesis [55, 85, 84, 114, 86, 87], and experiments to disprove it [54, 140, 63, 64, 10].

2.2.2 Ehrlich's side-chain theory

In 1897, Paul Ehrlich [37] described the interaction of diphteria toxin and antitoxin and how to measure it. He suggested that a unique stereochemical relationship between the active sites of an antigen and its antibody was responsible for immunological specificity. He also introduced the concepts of "affinity" and of "functional domains" on the antibody molecule. At the same time, he presented his side-chain theory which relied on

intracellular digestion, like Elie Metchnikoff's earlier phagocytic theory of immunity. He postulated that every cell capable of synthesizing antibodies had on its surface "side-chains" of different specificities. The antigen would react with its specific side-chain and be taken up by the cell. The side-chain could then perform its function anew or else it could be regenerated. However, in presence of large or repeated amounts of antigen, the cell would overproduce the specific side-chains and the excess would be released as antitoxins into the blood. This theory already had all the requirements of a natural selection theory. When it was proposed, only a limited number of antibodies were known, so it seemed possible for a cell to have side-chains of all the required specificities on its surface. But within a few years, numerous other antibodies were found, discrediting the side-chain theory. Figure 2.1 is one of the diagrams [38] that Ehrlich presented to illustrate his side-chain theory.

2.2.3 Instruction theories

The instruction theories were based on the assumption that the amount of information required for an antibody repertoire is so large that it cannot be contained inside the body and hence must come from the outside. The first theories of that kind were the direct template theories.

Direct template theories

In 1909, Oskar Bail and his co-workers [6] suggested that the antigen is not eliminated after interaction with its specific antibody. It would then free the specific antibody to pursue its function which is to bind to "natural antibodies" of the normal blood and thereby imprint these with its specificity. Theories built on the same basic idea were presented by other immunologists [5, 4, 138, 99, 100] until the beginning of the 1930's. It was then known that proteins were made of some kind of assembly of 20-odd types of so-called

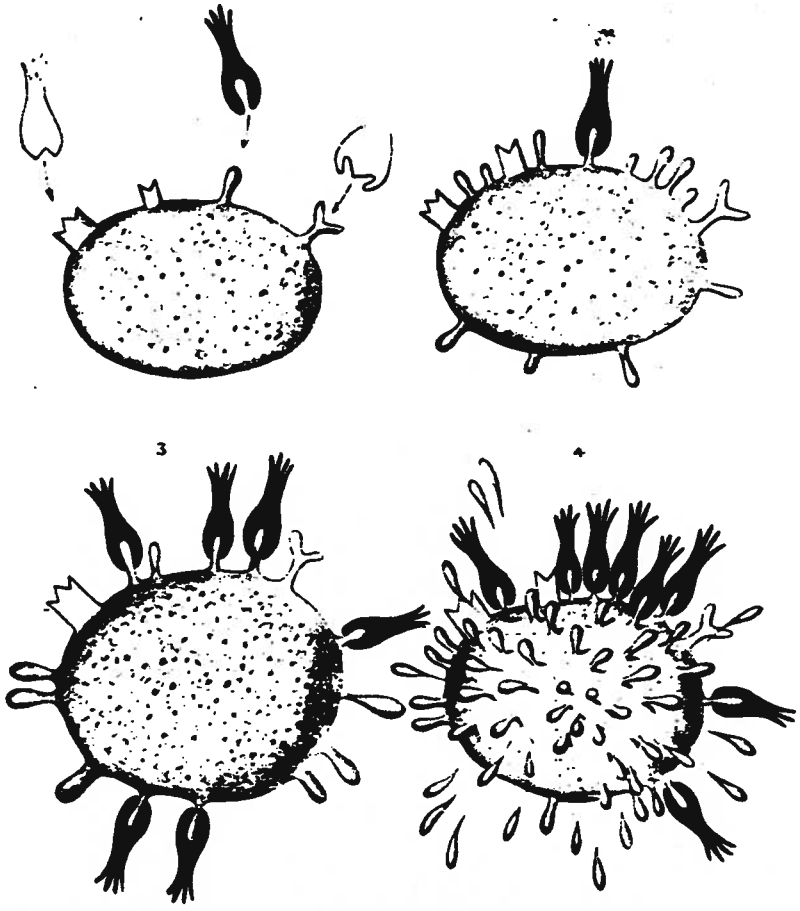

Figure 2.1: Ehrlich's side-chain theory: "the antitoxines represent nothing more than side-chains reproduced in excess during regeneration and therefore pushed off from the protoplasma, and so coming to exist in a free state". (Scanned from [38].)

amino acids and that antibodies were globular proteins. A new theory was then proposed almost simultaneously by various immunologists [141, 94, 3], but most conclusively by Breinl and Haurowitz [16, 53]. They suggested that the antigen travels to the site of protein synthesis and acts as a template for building up the nascent antibody molecule; further, through an unknown mechanism, the stereochemical structure of the antigenic site determines a unique amino acid sequence, which results in the complementarity and specificity of the interaction between the antibody and the antigen.

In 1940, the interaction of complementary three-dimensional configurations of atoms was formally shown by Pauling and his students to be responsible for the specificity of the antibody-hapten interaction, as Paul Ehrlich had suggested long ago. Their binding energy was apparently a combination of ionic, hydrogen-bonding and van der Waals interactions. Following this, Pauling [106] suggested that the specificity of an antibody must be due to its unique tertiary structure which is determined by a unique folding of its peptide chain. The antigen would serve as a template only at the time of coiling of the nascent polypeptide chain of the antibody molecule and the resulting configuration would be stabilized by familiar interatomic bonds. Figure 2.2 shows one of the diagrams that Pauling used to illustrate his "theory of the structure and process of formation of antibodies". This theory was extended by Karush [73], who proposed that the unique folding of the peptide chain was stabilized by multiple disulfide bridge cross-linkages, which with their different extents would produce the variety of antibodies. He also argued that the linearity of any template that determines a primary amino acid sequence is mandatory, which would invalidate the Breinl-Haurowitz theory.

One major problem with the direct template theories, not even mentioned by its proponents, was that they could not account for a greater and faster antibody response due to a second booster encounter with an antigen.

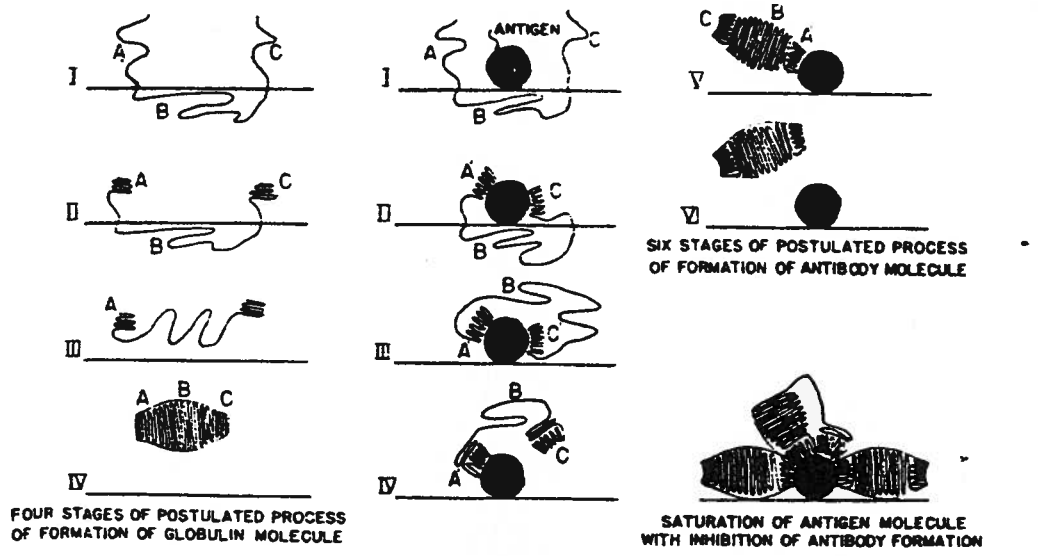

Figure 2.2: Diagram of Pauling's theory of antibody formation, representing the four stages in the process of formation of a molecule of normal serum globulin, the six stages in the process of formation of an antibody molecule as the result of interaction of the globulin polypeptide chain with an antigen molecule and the antigen molecule surrounded by attached molecules or parts of molecules and thus inhibited from further antibody formation. (Scanned from [106].)

Indirect template theories

In 1941, Burnet [20] presented his "adaptive enzyme theory", which later led to the indirect template theories. At the time it was thought that all proteins were broken down and synthesized by special proteinase enzymes. Some experiments had also suggested that "adaptive" enzymes could appear in response to special changes or needs of the bacterial organism. Burnet suggested that the antigen would reach the cells of the reticuloendothelial system and there get in contact with the local proteinases. Dissolution of the antigen molecule would be accompanied by an adaptive transformation of those enzymes, which could then synthesize a globulin molecule specific to the antigen. The information carried by those enzymes would be passed on to their daughter cells, thus enlarging the number of antigen-specific enzymes, and this would explain the stronger response obtained with large or repeated antigen injections. Furthermore, the theory would account for the recent observation that the affinity of the antibodies would improve after these booster injections, since the enzyme would adapt at each injection and thus become more specific. This was a net advantage over the direct template theories.

Towards the end of the 1940's, adaptive enzymes were no longer popular. It was known that a "genome" of undetermined composition ruled proteins synthesis. So in 1949, Burnet and Fenner [23] presented a new indirect template theory that would endow the antigen the ability to imprint its specificity on the genome, thus transfering it to the antibody. The information would persist in the genome of a cell and be passed on to the daughter cells, which would explain an enhanced secondary response. Reexposure to antigen would also improve the quality of the genocopy. Some experiments on grafts had just suggested the rationale of an immune system able to distinguish between self and non-self. So Burnet and Fenner postulated that during development, the body elements get "self-markers" which protect them from future attacks from the immune system.

By 1957, it was known that DNA is the carrier of genetic information. Schweet and Owen [123] presented then their "template-inducer" theory in which they proposed that the antigen first changes the DNA of the globulin gene, giving somatically heritable information for the generation of a new RNA template to produce cells "primed" for specific antibody formation. The antigen would furthermore act as an inducer on those cells, stimulating the formation of many templates and enhanced antibody production.

2.2.4 Selection theories

The first of the selection theories was presented by the physicist Jordan [72] in 1940. In his quantum-mechanical resonance theory, he suggested that the antigen, during its partial digestion inside the body, is divided into fragments which *select* and interact with *natural* molecules, their specific antibodies. Molecules having identical groups get attracted to each other through quantum mechanical resonance and this leads to autocatalytic reproduction of the antibody molecule. The structure of the daughter molecules can be altered during reproduction, which would account for the graded cross-reactions observed by Landsteiner [81]. His theory was attacked by Pauling [105] who stated that resonance attractions were more likely between complementary molecules than between identical molecules.

In 1955, Niels Jerne [66] presented his natural selection theory. He postulated that the antigen, after interaction of its specific determinants with their *natural* specific antibodies, would carry the antibodies to specialized cells able to produce them. The *selected* antibodies would induce synthesis of a specific RNA or change the structure of the existing RNA to allow for synthesis of more specific antibody molecules. The theory accounted for larger booster response and for improved affinity. Jerne also explained immunological tolerance by suggesting that during embryogenesis the first natural antibodies made against self-antigens are integrated by the body tissues and thus become

absent as stimuli for further autoantibody formation. But the theory had a problem that Talmage [136] pointed out, due to the so-called "central dogma" of genetics later formalized by Francis Crick [25]. It was that the instruction for protein structure came from the DNA via the RNA to the protein and stayed there. This process could not be reversed. This meant that none of the antigen or the antibody could directly inform the DNA on the production of a specific antibody. It could at most trigger an existing program. Ehrlich had already suggested this idea and Burnet and others returned to it.

The **clonal selection theory of antibodies** was first presented in a general way by Talmage [136] and Burnet [21], and then more extensively by Burnet [22], Talmage [137] and Ledeberg [82]. In his book published in 1959, Burnet suggested that *natural* receptors or antibodies possessing a unique specificity are situated on the surface of every lymphoid cell. The antigen interacts *selectively* with its specific receptor and triggers the differentiation of antibody production and proliferation to form daughter cells which constitute a clone of the same specificity. This already accounted for enhanced secondary responses and changes in quality of the antibody, the latter possibly improved by somatic mutations. Furthermore, Burnet explained self and acquired tolerance by postulating that clonal precursor cells might be especially susceptible to lethal action of their respective antigens early in ontogeny, which eliminates them from the repertoire. The antigen being "sequestered" and thus being absent for clonal elimination, as well as somatic mutations were suggested as the precursors events of auto-immune diseases.

Talmage theorized further about the roles of antigen selection and antigen-induced cell differentiation. He studied particularly the specificity and the size of the antibody repertoire and proposed that a heterogeneous immune system might show a greater specificity for an antigen than any of its constituent antibodies, because a distinct specificity would be displayed by each set of cross-reacting antibodies. Lederberg examined the genetic inferences of the clonal selection theory and suggested that somatic mutations continue

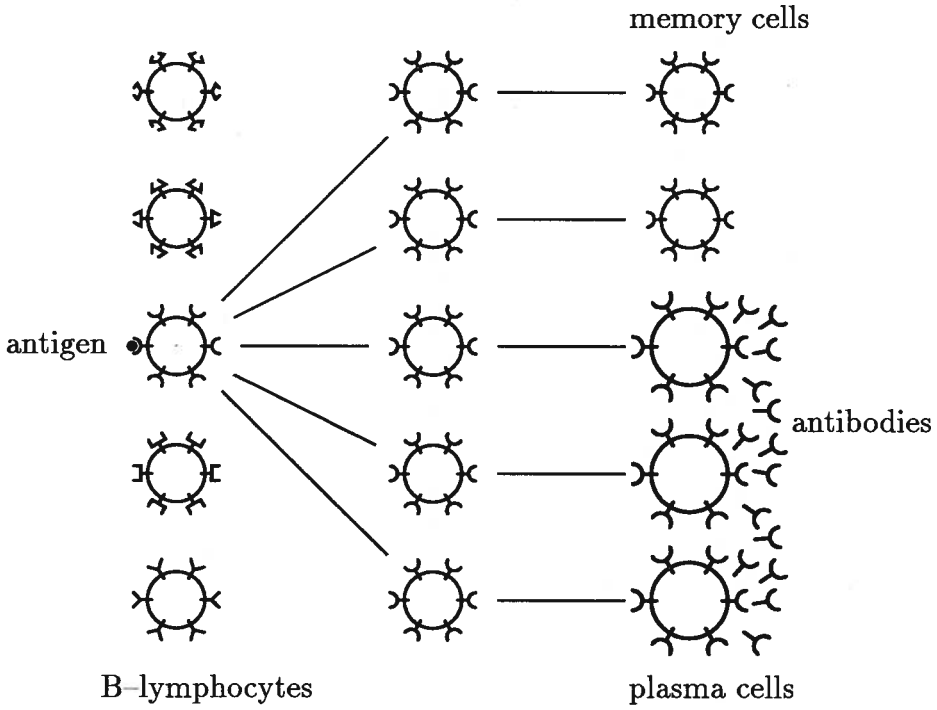

Figure 2.3: The clonal selection theory of antibodies: antigen binds to a receptor on the surface of a lymphocyte bearing the right specificity and triggers its proliferation into a clone of lymphocytes of the same specificity; some of the daughter cells enlarge to become plasma cells that secrete antibodies of the same specificity, others are the memory cells ready to respond to further stimulation by antigens of the same specificity.

to occur after fetal life [1]. Soon, the clonal selection theory became widely accepted in its general lines and is now considered to be the cornerstone of immunology.

[1]This idea of the somatic generation of antibody diversity would be the center of another important debate between supporters of the germ-line and the somatic theories, which would be concluded by both sides making concessions.

Chapter 3

Immune network theories and models

3.1 Jerne's network hypothesis

In 1974, Jerne [68] reviewed the pertinent immunological data that was available and postulated that the immune system functions as a network with a complexity comparable to that of the nervous system. To put his postulate in its context, let us make a summary of some basic facts of immunology which were generally accepted at the time.

It was known that the antibody is a protein molecule called immmunoglobulin (Ig) which is made up of four polypeptide chains: two identical light chains and two identical heavy chains, joined together by various disulfide bonds. The domain structure of the Ig molecule is illustrated in figure 3.4 borrowed from [50]. Both types of chains contain "constant" and "variable" regions. The five classes of immunoglobulins: IgG, IgM, IgA, IgD and IgE are determined by the constant regions of the heavy chains. The great diversity of antibodies within each class comes from the differences among the variable regions (V regions). These constitute the antibody combining sites (paratopes) which can react with the antigenic determinant (epitope). The importance of cells in immunity had finally been recognized, leading to the discovery that antibodies are actually produced by a class of white blood cells, the lymphocytes. There are two classes of these, namely the B cells originating from the bone marrow and the T cells derived from the thymus. Both possess specific receptors capable of recognizing antigens. The receptors of the B cells are immunoglobulins. Originally, the B cells are in their resting state and secrete

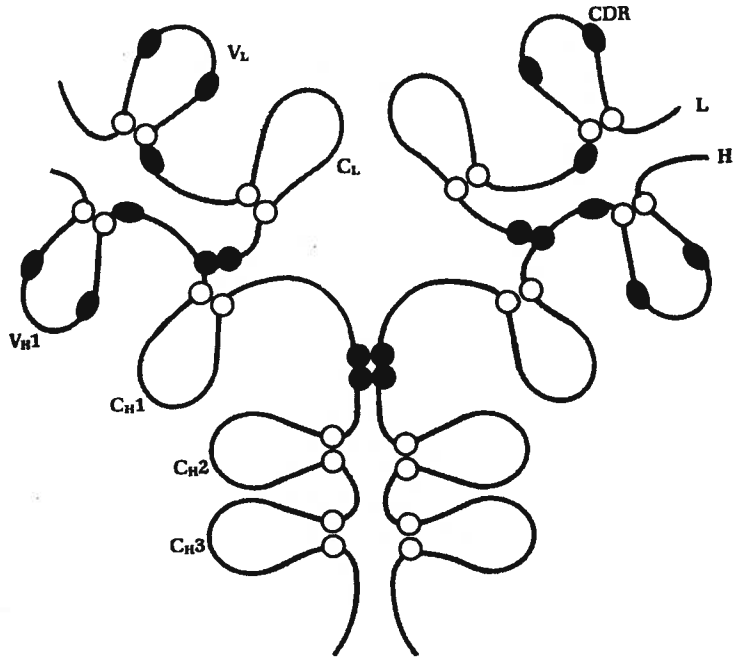

Figure 3.4: The domain structure of an Ig molecule. L and H represent the light and heavy chains respectively. C_L and V_L are the "constant" and "variable" regions of the light chain. C_H1, C_H2, C_H3 are constant regions of the heavy chains. V_H1 is the variable region of the heavy chain. CDR's are the complementarity determining regions, that is those regions that form the antigen-binding sites. (Scanned from [50].)

only a relatively small number of antibodies (mainly IgM) which they also display on their surface as "receptors". When the epitopes of antigens interact with paratopes of receptors, the B cells become either stimulated or suppressed. If they become stimulated, they begin to reproduce into a clone of cells all possessing the same specificity. Some of them enlarge and become "plasma" cells which begin to secrete large amounts of antibodies (mainly IgG) into the blood. The remainder reverts to the resting state and become "memory" cells, ready to respond if the antigen reappears. If the B cells become suppressed, they are no longer capable of being stimulated. This can happen, for example, with very high injections of antigens (high zone tolerance) or low injections below the threshold for stimulation (low zone tolerance).

The basis of Jerne's network hypothesis was that antibody molecules can recognize not only antigens but also other antibody molecules. Antibodies formed against the antigenic determinants of other antibodies had been discovered experimentally almost simultaneously by Jacques Oudin and Mauricette Michel [101] in France and by Philip Gell and Andrew Kelus [47] in England. They were called "anti-clone" antibodies by Gell and Kelus. The term "idiotype" was used by Oudin [101] to denote the antigenic specificities of antibodies produced by an individual or a group of individuals in response to a given antigen: that is, an idiotype is the set of epitopes on the variable regions of a set of antibodies. Jerne named each single idiotypic epitope an "idiotope". He thought that antibody molecules possess combining sites called idiotopes which can be recognized by paratopes of other antibody molecules, and viewed the immune system as a network involving stimulatory and suppressive interactions between idiotopes and paratopes. The system was normally maintained in equilibrium by the dominating suppressive interactions. Injection of an antigen into the system would disturb that equilibrium by giving more stimulation to clones with certain idiotopes. These would in turn stimulate other clones with anti-idiotopes specific to the idiotopes, and so on, leading to a chain reaction

that could potentially spread throughout the whole network. If the system then settled into a new equilibrium, this would represent immunological memory. Jerne [67] says: "I am convinced that the description of the immune system as a functional network of lymphocytes and antibody molecules is essential to its understanding, and that the network as a whole functions in a way that is peculiar to and characteristic of the internal interactions of the elements of the immune system itself: it displays what I call an eigen-behaviour." This was then almost the first time [129] that the production of antibodies against self-antigen was suggested to be normal rather than exceptional.

It is often forgotten that the main features of this hypothesis had already been put forward implicitly by Ehrlich in his side-chain theory in 1897. Furthermore, as Kossel [79] in Germany and Camus and Gley [24] in France had claimed having discovered anti-antibodies in 1898-9, Bordet [13, 14] on one side and Ehrlich and Morgenroth [39, 40, 41] on the other side had performed more experiments and speculated further on this discovery. They had been imitated by Alexandre Besredka [11], Pfeiffer and Friedberger [112], August von Wasserman [146], Hans Sachs [122] and others. By 1905, it had been realized that data had been misinterpreted, and so had died the first series of anti-antibodies theories. But even though they had been built on misinterpreted experiments, conceptually they had not differed much from what would be presented some seventy years later by Jerne.

3.2 First mathematical models of the network hypothesis

3.2.1 Introduction

The first models of Jerne's hypothesis [69] appeared during the year that followed his 1974 paper. Before examining them, it is appropriate to mention their precursor which was one of the first mathematical model proposed to describe clonal selection and antibody

production, presented by Bell [7, 8, 9] in 1970-1971. His model can be understood from the point of view of a network, because the absence of any interactions between cells or antibodies is equivalent to a network with zero strength of connections. Bell made no distinction between different classes of antibodies and ignored the T cell activity. He assumed that only a few lymphocytes can respond to an antigen to different degrees depending on the average properties of the association constants of their receptors for the antigen. Upon stimulation by an antigen, the lymphocyte can become paralyzed, or divide into two proliferating lymphocytes, or divide into either two plasma cells or a plasma cell and a memory cell, this latter being a new target cell for antigen. Antibodies secreted by plasma cells help to eliminate the antigen. This limited model embodied into a set of differential equations, accounted reasonably well for experimental antibody responses, for high and low zone tolerance and for changes in the affinity between a lymphocyte and its stimulating antigen.

3.2.2 Richter's model

The first model subsequent to Jerne's hypothesis including network interactions was presented by Richter [117] in 1975 and includes excitatory, suppressive and inhibitory interactions between the variable regions (V-regions) of the lymphocytes and of the antibodies. It ignores the differences between B and T cells and between antibody classes. An antigen stimulates a specific set of idiotypic antibodies which, in turn, stimulates a set of anti-idiotypic antibodies and so on. Negative feedback between idiotypic populations is assumed to occur to limit the growth of the populations. This negative feedback can happen only subsequent to the presence of stimulation by antigen. The set of differential equations representing the model can simulate low zone tolerance, a normal immune response and high zone tolerance, as well as immunological memory. It simulates several important features of immune responses and leads to testable expectations.

3.2.3 Hoffmann's model

The second model to appear in 1975 was due to Hoffmann [56]. It will be examined in more detail since we will use its latest version as a basis for our investigations. A new feature of this model is that it gets rid of the distinction between paratopes and idiotopes by assuming that the same surface pattern that can recognize antigen can itself be recognized by the variable regions of other lymphocytes and antibodies. More precisely, it is assumed that by averaging the three-dimensional shapes of all the V regions that are recognized by an antigen, one should obtain a shape which is somewhat complementary to the shape of the epitope of the antigen. The system is therefore approximated by two sets of lymphocytes, the "positive" set that interacts with the antigen, and the "negative" set that interacts with the positive set. The symmetrical interaction between the two sets is postulated to happen via cross-linking of receptors. Thus the positive set can stimulate the negative set to proliferate and vice versa.

The difference between T and B cells is taken into account. T cells secrete monovalent specific factors. Specific T cell factors produced by one set can block the receptors of the T and B cells of the other set and thus inhibit interactions of the two sets. B cells of one set secrete antibody molecules which can kill, in association with complement, T and B cells of the other set. This is called "complement-mediated cytotoxicity". One antibody molecule of the IgM type or two of the IgG type bound to the cell surface are required for this process. B cells switch from producing mainly IgM molecules to producing mainly IgG molecules during the course of an immune response. Effector cells (e.g macrophages), to which the constant region of the antigen specific T cell factor can bind, were postulated to be able to kill T and B cells of the complementary specificity. This was called "indirect cell-mediated cytotoxicity".

Four stable states of the system are postulated to be: a virgin state where populations

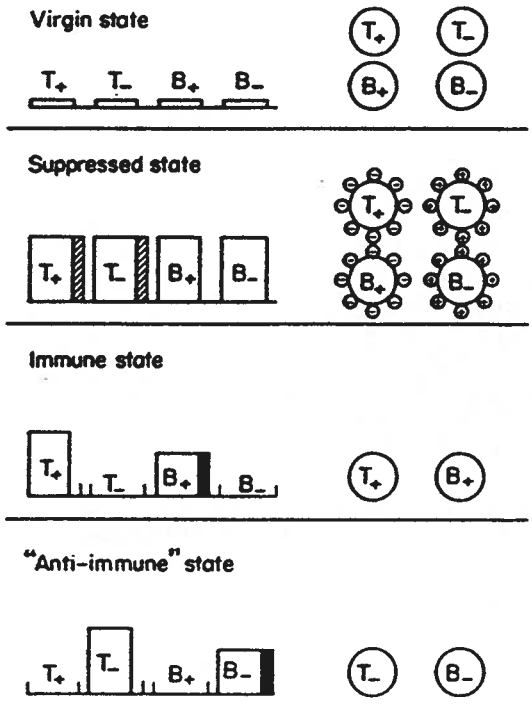

Figure 3.5: The stable states of Hoffmann's model. (Scanned from [56].)

of both sets are low enough that there is not a significant amount of stimulation but there is killing (linear), an immune state where the populations of positive cells is high and the immune effector function kills the negative population, a suppressed state where populations of T cells of both sets is high and where all stimulation is inhibited by a high level of monovalent factors, and finally an anti-immune state which is the converse of the immune state. Figure 3.5 illustrates the stable states of the model.

Hoffmann presents a relatively complex set of differential equations that takes into account the theory described above and generates the steady-states. For simplification, he assumes that the concentration of antibodies is proportional to the concentration of the cells that produce them, thus allowing the use of only one variable to model both of these. The difference between B and T cells is taken into account by different functions

and kinetic constants. The antigen, by modulating the dynamics of its specific cells, takes care of the difference between cells of each specificity. The model also includes the two types of killing, cell and complement mediated, and for this latter more specifically the difference between IgM and IgG killing. The switch from IgM production to IgG is also included. The theory also deals with with several other features of the immune system such as self-tolerance, and suggests experiments.

3.2.4 Adam-Weiler's model

During the same year of 1975, Adam and Weiler [2] presented a model which aimed at explaining the generation of antibody diversity and of a large set of different lymphocyte clones in early ontogeny.

3.3 Later models

We have just examined the very first models of Jerne's network hypothesis that appeared during the following year that is was proposed. Those first modelers pursued their investigations to different extents. Adam [1] published one more paper in 1978. Richter [118, 119] published two other papers in 1978. Hoffmann still continues his research in the field. His work will be examined in more details in the next chapter.

Many other researchers have investigated the network hypothesis. Some of the most important ones are mentioned here.

Kaufman, Urbain and Thomas [76] produced a model in 1985 which is based on a logical analysis of the immune response, using boolean variables. In 1987, the continuation of that paper is presented by Kauman and Thomas [75]. They analyse the same model with continuous variables this time. Kaufman, Urbain and Thomas have published more after that and are still continuing their studies.

De Boer [27, 28] presents his first model in 1988. His model differs from the previous ones in that he does not consider suppression to be the important mechanism for maintaining homeostasis in the virgin state. He is continuing his investigations, in conjunction with others [29, 30, 31, 32, 33].

Here, it is certainly not the aim to do an extensive review of all the literature on studies of immune networks. But the names of some of the other researchers in the field that have persevered over the years are given here for reference (in alphabetical order) with some of their papers: Avidan Neumann [95, 96], Alan Perelson [110, 109], Lee Segel [124], Dietrich Stauffer [131, 134, 133, 132, 131, 134], Francisco Varela with Antonio Coutinho and John Stewart [142, 135], Gerard Weisbuch [149, 151, 150]. It should be well understood that this is not an exhaustive list.

3.4 More on Hoffmann's model

Here, we will follow the development of Hoffmann's model in the following years. The reason for doing so is that it will help the reader to understand gradually the apparently complex form of the latest version, the N-dimensional network model presented in §3.5.

A symmetrical two-dimensional model: the "plus-minus" model

In 1979, Hoffmann [57] presented the "plus-minus" model that was aimed at a better understanding of the behaviour of the immune system near the steady-states rather than the more complex events that occur during the switching between them. But his postulates were essentially the same. He was also more interested in a qualitative rather than quantitative understanding of the steady-states. So he made several further simplifications to the ones already made in his previous model. First, he made no difference between the

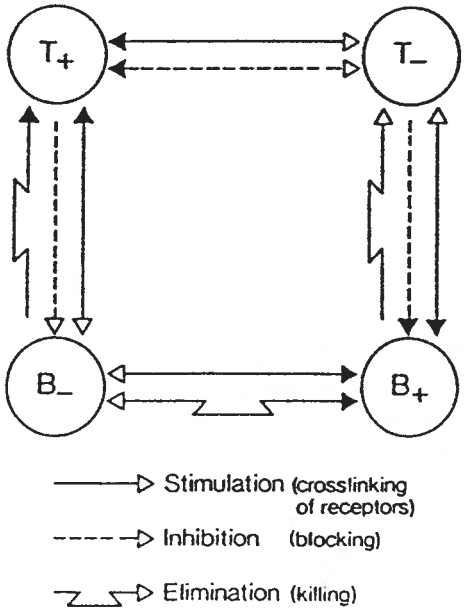

Figure 3.6: The interactions between cells in the plus-minus model. (Scanned from [57].)

T and B cells in the mathematical model, on the basis that at the level of approximation of the model the same selective forces act on T_+ cells and B_+ cells, and similarly for T_- and B_- cells, where T_+, B_+, T_- and B_- cells represent T and B cells of the positive and of the negative sets. Figure 3.6 illustrates the interactions between cells in the "plus-minus" model. Accessory cells, previously termed effector cells, were assumed to be involved in the switching between the steady-states but not in the steady-states themselves. They were thus no longer considered in the mathematical model; similarly, the antigen did not enter the model. The concentration of specific T cell factors was chosen to be proportional to the product of the concentration of the cells that produce them and of the concentration of cells that specifically stimulate those cells. And once again the concentration of antibodies was treated as being simply proportional to the concentration of cells that produce antibodies.

The "plus-minus" mathematical model is presented here since we will be using a

version which is simply its extension to N dimensions. If one denotes by x_+ and x_- the concentrations of the positive and negative sets respectively, the differential equations governing their dynamics are of the form:

$$\frac{dx_+}{dt} = S + k_1\, x_+\, x_-\, e_1 + k_2\, x_+\, x_-\, e_2 + k_3\, x_+\, (x_-)^2\, e_3 - k_4\, x_+ \tag{3.1}$$

$$\frac{dx_-}{dt} = S + k_1\, x_+\, x_-\, e_1 + k_2\, x_+\, x_-\, e_2 + k_3\, (x_+)^2\, x_-\, e_3 - k_4\, x_- \tag{3.2}$$

$$\text{where} \quad e_q = \frac{1}{1 + \left(\dfrac{x_+\, x_-}{{c_q}^2}\right)^{n_q}} \qquad q = 1, 2, 3. \tag{3.3}$$

The first term S models the constant non-specific influx of cells into the system. The second term models the mutual stimulation of the cells. The third term denotes killing by killer T-cells and/or by IgM plus complement. It is linear in the concentration of cells of the opposite specificity and has a rate constant equal to k_2. The fourth term models antibody-dependent cellular cytoxicity and/or killing by IgG plus complement. It has a quadratic dependence on the concentration of cells of the opposite specificity and has a rate constant equal to k_3. The last term models the natural non-specific death and has a rate constant equal to k_4.

The e_q terms model the specific inhibition of stimulation and killing by specific T-cell factors (figure 3.7). The interaction between positive and negative cells and antibodies are assumed to be inhibitable by both positive and negative specific T cell factors. The c_q are constants that specify the threshold values of the product $x_+\, x_-$ at which the inhibition becomes effective, and the n_q are constants that determine the sharpness of the thresholds. The product of the concentrations is used (rather than, say, the sum) because the concentration of T cell factors depends on the amount of stimulation between positive and negative cells. The fact that IgM production is more important in the virgin state than IgG production and vice versa in the immune state is incorporated in the model by choosing different values for the C_q and the n_q.

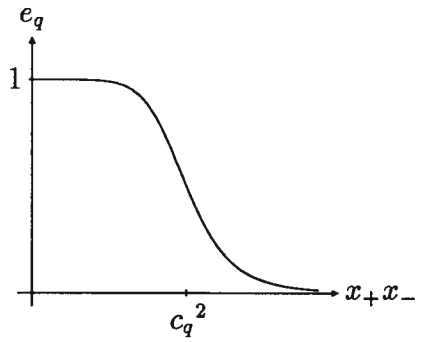

Figure 3.7: The e_q function for n_q=7.

Hoffmann finds parameters that lead to the postulated steady-states and performs some stability analysis on them. He finds that the steady-states are all attractors. Figure 3.8 is a phase-plane diagram of Hoffmann's symmetrical plus-minus model. He also studies the model when only one of the two IgM or IgG killing terms is present and finds that one does not then get all the steady-states or that they are then not all stable.

It should be noted that in a paper published in 1980, Hoffmann [58] reviews the experimental findings which support the idea of symmetric[1] interaction between idiotypes and anti-idiotypes of both T and B cells, cross-linking as the mechanism of interaction and the existence of specific T-cell factors.

Hoffmann's next aim is to find the *minimal* model that leads to the postulated set of stable-states. Gunther and Hoffmann [51] find that a minimal model does not require a positive stimulatory term (the term with rate constant equal to k_1), but only the negative suppressive ones.

[1] In 1984, Jerne [70] reviewed experimental data on idiotypic interactions and came to the conclusion that they are indeed symmetrical.

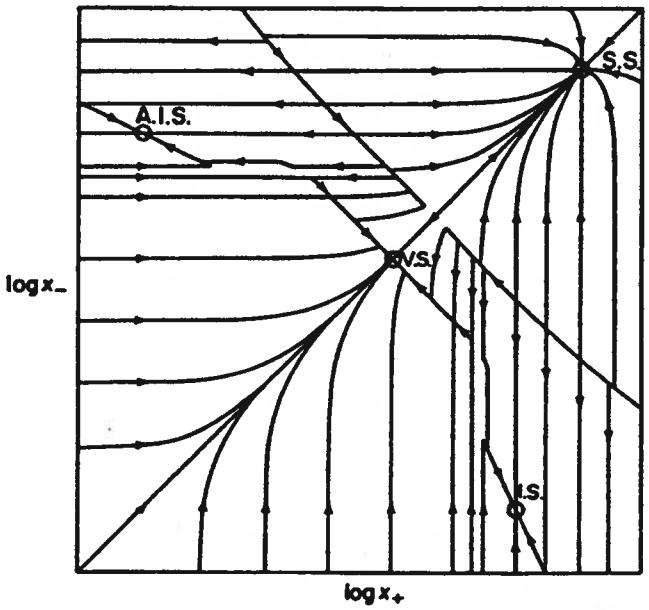

Figure 3.8: Phase-plane representation of Hoffmann's plus-minus model. The virgin, immune, anti-immune and suppressed states are labeled V.S., I.S., A.I.S. and S.S. respectively. All four stable states are attractors. Parameter values are $k_1 = 0.1$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0.01$, $S = 1$, $c_1 = 10$, $c_2 = 3$, $c_3 = 0.3$, $n_1 = 1$, $n_2 = 2$, $n_3 = 2$. (Scanned from [57].)

3.5 Hoffmann's N-dimensional model

3.5.1 The model

A first attempt to generalize the two dimensional "plus-minus" model to N dimensions appears in 1988, presented by Hoffmann et al. [60]. A second improved version by Hoffmann et al. [61] appears later, which gets rid of an inconsistency with regard to the symmetry of the interaction, overlooked in the first version, but brings unnecessary complexity to the model. The most recent and simplest version is achieved by Mathewson and Hoffmann [90] and is presented here, since it will used as a basis for further investigations.

One denotes the population of a clone i as x_i. For a network of N clones, the differential equation governing the dynamics of clone i has the form:

$$\frac{dx_i}{dt} = S + k_1\, x_i\, U_{1i}\, e_{1i} - k_2\, x_i\, U_{2i}\, e_{2i} - k_3\, x_i\, (U_{3i})^2\, e_{3i} - k_4\, x_i \qquad i = 1, N \qquad (3.4)$$

$$\text{where } e_{qi} = \frac{1}{1 + \left(\dfrac{W_i'}{C_q}\right)^{n_q}} \qquad q = 1, 2, 3. \qquad (3.5)$$

One can see that equation (3.4) now replaces equations (3.2) and (3.1), with various values of the index i now representing clones of various specificities (in the two dimensional case, the plus and minus specificities). The only difference is the term U_{qi} (where $q = 1, 2, 3$) instead of the concentration of clones of the other specificity (plus or minus). We also now have equation (3.5) instead of equation (3.3) with the term W_i' replacing the previous product of the concentrations of plus and minus clones; the term C_q is used instead of ${c_q}^2$ for simplicity.

Consider the term W_i'. To be consistent with the two-dimensional model, it must represent a measure of the concentration of T cell factors. Before giving its value in the N-dimensional case, let us first introduce other quantities that are relevant in an N-dimensional network of interacting clones. Clones in the network will interact with each

other with different strengths of interaction (affinities), depending on their specificities. One can denote by K_{kj} the affinity between clones k and j. Studies of Hoffmann's N-dimensional network have been done up to now only with Boolean affinities, that is $K_{ij} = 0$ for no interaction between clones i and j, and $K_{ij} = 1$ for an interaction between clones i and j. The implementation of non-Boolean affinities in Hoffmann's N-dimensional model (as will be seen later) is one of the novelties of the work performed for this thesis. Another useful quantity is the field of a clone k, given by:

$$Y_k = \sum_{j=1}^{N} K_{kj}\, x_j \qquad\qquad k = 1, N. \tag{3.6}$$

It is the sum of the concentrations of all clones j that interact with clone k, weighted by the strength of each interaction. A measure of the T cell concentration of a clone i is proposed as:

$$W_i = \sum_{k=1}^{N} K_{ik}\, x_k\, Y_k. \tag{3.7}$$

One can see that this expression is a generalization of the two-dimensional one in the sense that it is proportional to the product of the concentrations of two clones for $N = 2$. Hoffmann et al. [61] found that this formulation was creating some problems in that, in their numerical simulations, all the clones tended towards the suppressed state. So they added a refinement to this formulation. They postulated that an enzyme controls the total amount of T cell factors. It keeps the total amount of T cell factors constant, more precisely equal to N times the average of the threshold constants C_q. This can be done by first calculating W_i, then multiplying it by a factor equal to the average of the two threshold constants divided by the average of all the T cell factor concentrations, so that the new expression for the T cell factor concentration is given by:

$$W_i' = W_i \frac{(C_2 + C_3)/2}{\dfrac{1}{N}\displaystyle\sum_{i=1}^{N} W_i}. \tag{3.8}$$

The other term we need to look at is the U_{qi}. It is called the effective field and is modelled by:

$$U_{qi} = \sum_{j=1}^{N} K_{ij}\, e_{qj}\, x_j \qquad\qquad q = 1, 2, 3. \tag{3.9}$$

The summation takes into account the fact that clone i potentially receives stimulation from all other clones j in the network. Each clone j can stimulate clone i more or less efficiently depending on its specificity, which is taken into account by the affinity K_{ij} between clones i and j. This stimulation can be inhibited if the level of T cell factors of clone j is too high, reflected by e_{qj} in equation (3.9), and/or if the level of T cell factors of clone i is too high, reflected by e_{qi} in equation (3.4).

A last comment is that in their recent simulations of the N-dimensional network, Hoffmann and his co-workers have used effectivities e_{qi} with a very sharp threshold. They have found that the results were not significantly affected by this choice [91]. We have decided to keep the same choice in order to have results that can complement theirs. For sharp threshold, that is $n_q = \infty$, the effectivities become:

$$e_{qi} = \begin{cases} 1 & \text{if } W_i \leq C_q \\ 0 & \text{otherwise.} \end{cases} \tag{3.10}$$

3.5.2 The steady-states of the model

We will consider the minimal model for which $k_1 = 0$. The theory predicts three steady-states for the N-dimensional model: a virgin state where the production of IgM antibodies is dominant, an immune state where the producion of IgG antibodies becomes more important and a suppressed state where the production of both types of antibodies is inhibited. The state of a clone is determined by W_i', the amount of its T cell factors. Depending on the values of the threshold constants C_2 and C_3, there are two manners to obtain the predicted steady-states.

Case $C_2 > C_3$. A clone is in the virgin state if $C_3 < W_v' < C_2$. The steady-state is reached when:

$$\frac{dx_v}{dt} = S - k_2\, x_v\, U_{2v} - k_4\, x_v = 0. \tag{3.11}$$

The steady-state has the value:

$$x_v = \frac{S}{k_2\, U_{2v} + k_4}. \tag{3.12}$$

A clone is in the immune state if $W_v' < C_3 < C_2$. The steady-state is reached when:

$$\frac{dx_i}{dt} = S - k_2\, x_i\, U_{2i} - k_3\, x_i\, U_{3i}^2 - k_4\, x_i = 0. \tag{3.13}$$

The value of the steady-state is given by:

$$x_i = \frac{S}{k_2\, U_{2i} + k_3\, U_{3i}^2 + k_4}. \tag{3.14}$$

A clone is in the suppressed state if $C_3 < C_2 < W_v'$. The steady-state is reached when:

$$\frac{dx_s}{dt} = S - k_4\, x_s. \tag{3.15}$$

This yields:

$$x_s = \frac{S}{k_4}. \tag{3.16}$$

Case $C_2 < C_3$. A clone is in the virgin state if $W_v' < C_2 < C_3$. The steady-state is reached when:

$$\frac{dx_v}{dt} = S - k_2\, x_v\, U_{2v} - k_3\, x_v\, U_{3v}^2 - k_4\, x_v = 0. \tag{3.17}$$

The value of the steady-state is given by:

$$x_v = \frac{S}{k_2\, U_{2v} + k_3\, U_{3v}^2 + k_4}. \tag{3.18}$$

A clone in the immune state is characterized by $C_2 < W_i' < C_3$. Then the steady-state is reached when:

$$\frac{dx_i}{dt} = S - k_3\, x_i\, U_{3i}^2 - k_4\, x_i = 0. \tag{3.19}$$

The steady-state has the value of:

$$x_i = \frac{S}{k_3 U_{3i}^2 + k_4}. \tag{3.20}$$

The suppressed state happens if $C_2 < C_3 < W_i'$. Then the steady-state is reached when:

$$\frac{dx_s}{dt} = S - k_4 x_s. \tag{3.21}$$

This yields:

$$x_s = \frac{S}{k_4}. \tag{3.22}$$

At this moment, further analysis on the N-dimensional network model is still unpublished [90]. But it is available in the Master's thesis of D. Mathewson [89]. In further work that is presented in this thesis, only the case $C_2 < C_3$ will be considered.

Chapter 4

The affinity distribution

4.1 The problem of the reconstruction of the affinity distribution

Introduction

A great problem that theorists face when they try to model the immune network is the lack of experimental data available for the parameter values of their models. This concern seems to be most prominent in the determination of the distribution of the equilibrium binding constants (affinities). The affinity K between two substances A and B is defined such that for any reaction:

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}, \tag{4.23}$$

one has that:

$$K = \frac{[\mathrm{AB}]}{[\mathrm{A}]\,[\mathrm{B}]}. \tag{4.24}$$

It is known from experimental evidence that the distribution of the affinities between an "invading" antigen and the antibodies produced by the lymphocytes is heterogeneous and varies with time under various circumstances [107, 42, 43, 48]. Many attempts have been made to reconstruct the affinity distribution from the relatively few measured values, which are available only for a finite number of free hapten concentrations. One major drawback when one attempts to do so is that several solutions can fit the data.

Gaussian and Sips distributions

Pauling et al. [107] tried a Gaussian distribution in 1944, followed by Eisen and Karush [42] in 1949. The Sips distribution was proposed by Nisonoff and Pressman [97] in 1958 and by Karush [74] in 1962. The affinity mean value and the variance of those distributions were estimated from the experimental data by non-linear regression. The interpretation of affinites for a small range can be done using the Gaussian and the Sips distribution but has been found to be inappropriate in general. One good reason for this is that the affinity distribution is now known to be asymmetric and/or bimodal [152, 120].

Fourier and Stieltjes transforms

Other attempts have been made to reconstruct the affinity distribution without preassigning a specific distribution. An exact analytic inverse based on the Fourier transform was proposed by Bowman and Aladjem [15]. It was never used in practice because it requires measurement of the experimental data with an extreme precision over the whole range of free hapten concentrations. Its advantage though was to provide a unique solution in the analytical sense. A similar approach was used by Bruni et al. [17, 18] using the Stieltjes transform. They offered a way to interpolate and extrapolate the data to generate a complete binding curve for the inversion. The problem with this is that the reconstruction of the affinity distribution can be strongly influenced by the details and manner of this completion. Hsu [65] suggested state-variable reduced order procedures. His method was also never used in practice because it requires both the complete binding curve and its derivative.

Histograms and delta functions

Other approaches to the problem of the reconstruction of the affinity distribution include a method based on histograms by Werblin and Siskind [152] in 1972 and another one on delta functions by Erwin and Aladjem [44] in 1976. The problem with these two methods is that several histograms or delta distributions can be found to fit the limited data available.

Minimum cross-entropy

In 1991, Yee [154] proposed a method based on minimum cross-entropy. It is a highly non-linear inversion procedure that does not require any prior assumptions about the affinity distribution or of parts of it. It apparently provides a good recovery of the distribution from a very limited amount of data. At this point, it is not known whether this method has been implemented by experimentalists.

4.2 Choices of connectivity matrices

From the previous discussion of attempts to reconstruct the affinity distribution, it can be seen that no solid ground is available for immune network theorists to work on. Under this uncertainty, they have chosen diverse connectivity matrices (matrices of affinities between the idiotypes of the immune network) for the simulations of their models.

Experimental connectivity

Varela et al. [142] have chosen affinities derived from experimental findings. They have given the affinities values 0 or 1 taken from connectivity matrices based on cross-reactivity as measured by ELISA assays.

Random connectivity

Others such as Hoffmann [60], Parisi [102], De Boer [30] have used random and symmetrical connectivity matrices. Weisbuch [150] has used a random but asymmetrical connectivity matrix. The reasons for choosing a random connectivity matrix are well expressed by Hoffmann [59], and by Weisbuch [150]. As Hoffmann mentions, somatic mutation processes produce a considerable part of the immune system repertoire of V-regions (variable regions), which renders their exact knowledge impossible for a particular animal. This means that the immune system must be constructed in such a way that it is not important which V-regions are present, as long as a very diverse repertoire exists, such that for each idiotype there are some matching anti-idiotypes. He continues by stating that the actual V-regions are not completely random since they are produced by combinations and mutations of a finite set of germ line genes. It seems that in some strains, particular "germ line" antibodies (or idiotypes) occur reproducibly. But considering that diversity has developed under various somatic mechanisms and especially the fact that repertoires can be mixed (by crossing strains, for example) without ill effect, one may deduce that the stability of the network should not depend on a specific structure of the connectivity matrix. The reasons that Weisbuch mentions for choosing a random connectivity matrix complement the ones given by Hoffmann. He suggests that the use of a random connectivity matrix is related to not knowing the real connectivity matrix. He also mentions it as a "good choice" since one searches for general properties that should not rely in any critical way on a specific structure of the matrix which might vary from one animal to the other.

Bit-string complementarity

Some have decided to model the affinities using the bit-string approach [45, 32]. They associated the shape of the receptors with a binary string and used the degree of complementarity between the bitstrings to determine the affinity between the clones.

Shape-space concept

A recently popular approach is to use the concept of shape-space first introduced in an immunological context by Perelson and Oster [110] in 1979. The affinity between two receptors depends on factors such as the average geometric shape of the binding region and the intermolecular attractive forces (hydrogen bonding, electrostatic forces, van der Waals bonds and hydrophobic bonds). Perelson and Oster [110] have shown that a small number (five to ten) of measurements of such factors is enough to describe adequately the "generalized shape" of the binding region.

This approach was first used by Perelson and Segel [124] in 1988. Each idiotype is associated with a vector of real numbers that describes its generalized shape. They studied a one-dimensional shape-space, in which the variable x would describe, for example, the geometric shape of the binding region or its electric charge. Figure 4.9 is a diagram that they used to illustrate their "idealized" one-dimensional shape-space. Perelson and Segel numbered all the shapes from $-N$ to N and assigned maximum complementarity to the shapes $y = -x$ and x. They assumed that the smaller is $x - (-y) \equiv x + y$, the better is the fit of shape x with shape y. To simulate this, they used a Gaussian distribution centered around shape $-x$ to represent the affinities of shape x with the other shapes. More will be said about this in the next chapter.

In 1988, a shape-space for which no dimension needed to be specified was studied by Percus [108]. Then, a high-dimensional shape-space restricted to a Bethe lattice was

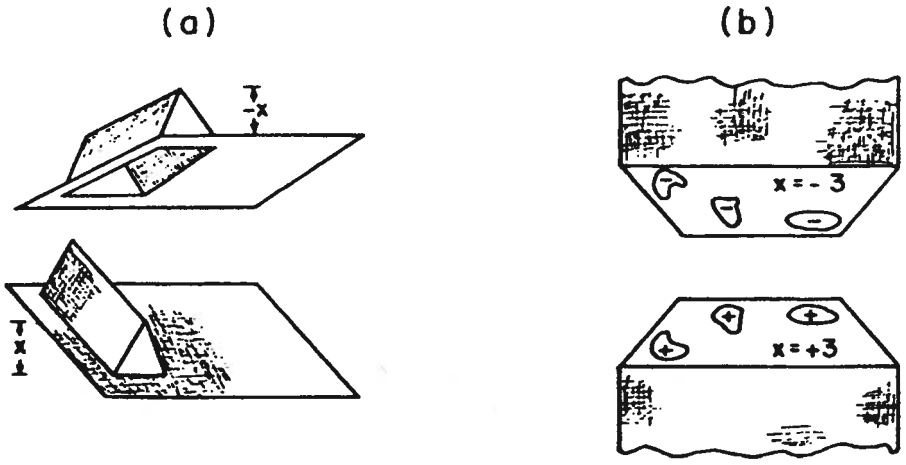

Figure 4.9: Two realizations of a one-dimensional shape-space. In figure (a), the shape is defined by the height of wedge-shaped epitopes, positive for protuberances and negative for indentations. In figure (b), the shape is defined as the charge of the homologously located patches. (Scanned from [124].)

used by Weisbuch et al. [151] in 1990. Studies of models of the immune network in two- and/or three-dimensional shape-space frameworks were done in 1990-1 by Weisbuch [150], Weinand [147, 148], and Stewart and Varela [135], and in 1992 by De Boer et al. [33]. Higher-dimensional studies have been done by Stauffer and Weisbuch [134]. De Boer et al. [34] have studied a model of the immune network for which there were only two possible states for the clones. This was an incentive for Stauffer and Sahimi [133, 132, 131] to use techniques developed from statistical physics (for example, for Ising models) to simulate a high-dimensional shape-space (up to ten dimensions) with a high-dimensional network (10^9 clones).

Chapter 5

Simulations in one-dimensional shape-space

5.1 Introduction

It was said previously that Hoffmann's models had only been studied with Boolean affinities. It was thus of interest to insert non-Boolean affinities in his model. The following studies in shape-space are not meant to be extensive, but rather to be short incursions which give an idea of the interesting behaviour that can arise from the implementation of an immune network model in shape-space. Since this had never been done with Hoffmann's model, it seemed quite relevant. But in fact, the original reason for choosing affinities of the shape-space was that an arbitrary choice of non-Boolean affinities was required to test numerically certain new analytical work about the "distance coefficient" which will be presented in the last chapter of this work. So, the reader is asked for his (her) indulgence with regard to its demands toward a more extensive study of such an interesting topic, that is the shape-space concept. He is reminded that this work is only made of brief incursions in several directions which call for further work.

In the following simulations, networks of dimensions approximately equal to 20 will be done since all previous work by others has been done for that dimension. Hoffmann's N-dimensional network model presented in 3.5 is implemented with the following fixed parameter values: $S = 10$, $k_2 = 1$, $k_3 = 0.01$, $k_4 = 1$ and others which will be specified when relevant.

5.2 The ± shape-space

5.2.1 Introduction

Hoffmann's N-dimensional model is first studied using the original idea of a shape-space proposed by Perelson and Segel (see §4.2). This shape-space will be referred to as the ± shape-space, since positive shapes are complementary to negative shapes in it. For the simulations, a network of dimension 21 is studied. First, the non-periodic case is considered, then the periodic case. The initial concentrations of the clones are chosen to be uniform or random. Some parameters are modified to observe the effects on the dynamics of the network.

5.2.2 The non-periodic ± shape-space

The clones are numbered from $-N$ to N and positioned on a finite one-dimensional shape-space axis accordingly to their similarity with each other (figure 5.10). The distance between two clones determines their similarity. Immediate neighbours on shape-space (like clones x and $x+1$) have the greatest similarity and are at an arbitrary distance of δ from one another. More distant clones are less similar. For example, clones $-N$ and N are the least similar. Clones of opposite sign (x and $-x$) are assigned maximum complementarity (maximum affinity). This is illustrated in figure 5.10 by dotted lines joining the clones of maximum complementarity. The complementarity of clone x with other clones situated symmetrically on each side of clone $-x$ decreases as these are being further apart from clone $-x$. For example, the complementarity of clone x with clones $-x-1$ or $-x+1$ is the second greatest. To model this, De Boer et al. [33] have used a "discrete" Gaussian centered at x. The same will be done here. Note that this shape-space representation allows for different combinations of similarity and complementarity between clones. For example, clone 0 has maximum complementarity with itself, as well

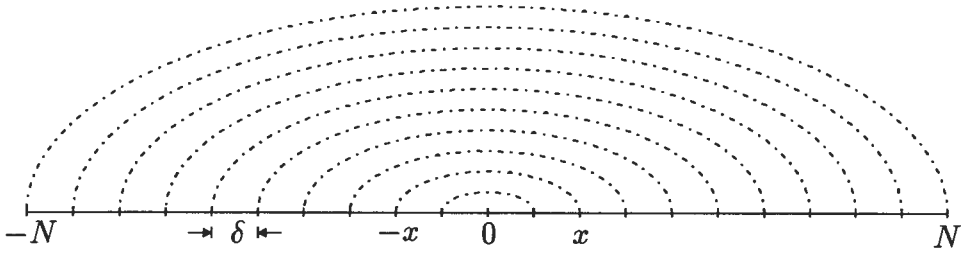

Figure 5.10: The non-periodic $\pm$ shape-space. Clones are numbered from $-N$ to N and positioned accordingly to their similarity with each other. Neighbouring clones are the most similar and are separated by an arbitrary distance δ. Dotted lines join the clones of maximum complementarity. Note that clone 0 has maximum complementarity with itself.

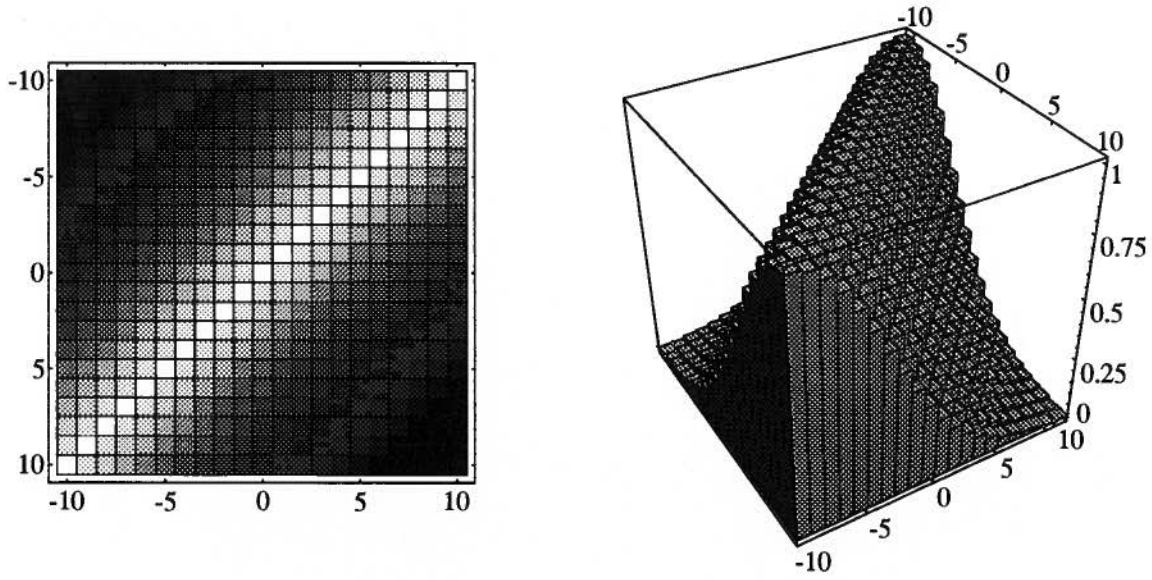

Figure 5.11: The matrix of affinities of a 21-dimensional network, for the non-periodic $\pm$ shape-space. The left and right pictures are two-dimensional and three-dimensional representations. Clones are numbered from -10 to 10 for both representations. In the two-dimensional representation, the shade of grey represents the value of the affinity with a scale going from white representing a maximum affinity of one and black representing a minimum affinity of zero. In the three-dimensional representation, the vertical axis represent the value of the affinity. Here, $\sigma = 6$ and *cut-off*$= 10^{-6}$.

as maximum similarity. But clone N has maximum complementarity with clone $-N$ and minimum similarity with itself.

The affinity between clones i and j is denoted by K_{ij}. It is of course equal to K_{ji}, by symmetry of the interaction between two clones. The $2N+1 \times 2N+1$ matrix of affinities shown for $N = 10$ in figure 5.11 is based on the following relationship:

$$K_{ij} = k\ (2\pi\sigma^2)^{-1/2} exp[-(i+j)^2\ \delta/2\sigma^2] \qquad i,j = -N, N. \qquad (5.25)$$

where k is a normalization constant, δ is the space between shapes on the one-dimensional shape-space axis, and σ^2 is the variance of the Gaussian. In all simulations, the normalization constant k was selected such that the maximum value taken by the affinity was one, that is $k = 2\pi\sigma^{1/2}$. The space δ was arbitrarily given the value 1. The only parameter that was varied was σ. To include non-periodic boundary conditions, the tails of the discrete Gaussians are truncated when they reach the boundaries of the shape-space axis. An additional parameter not shown here which appears in the Fortran program that generates the matrix of affinities is *cut-off*, the smallest possible value of the affinity; whenever $K_{ij} <$ *cut-off* it is set equal to zero. Such a simplification has been done by De Boer et al. [33]. The matrix of affinities has a band of Gaussian distributions that are centered on the upper-right to lower-left diagonal.

The non-periodic ± shape-space with random initial conditions

The system is studied with random initial clone populations, more specifically for $0 < x_i < 1$ at time t=0. The characteristic parameters of the Gaussian distribution of affinities are chosen to be: $\sigma = 4$ and *cut-off*$= 10^{-3}$. The other parameters of the model are $S = 10$, $k_2 = 1$, $k_3 = 0.01$, $k_4 = 1$, $C_2 = 0.6$ and $C_3 = 0.7$. The differential equation (3.4) and its dependents is integrated until time t=100, which allows for all clone populations to reach complete equilibrium.

Dynamics of the model Figures 5.12 to 5.17 show the dynamics for different variables of the model. Figure 5.12 represents the dynamics of the clone concentration x_i versus the time t. At time t=0, the clone concentrations are randomly distributed between 0 and 1. They increase to a value greater than the equilibrium value, decrease below it, and then increase again to reach it and settle there. At equilibrium, the clones have three distinct ranges of populations and complementary clones have the same population. Each pair of complementary clones has its own value of the population. Clones are in the three steady-states of equations (3.18), (3.20) and (3.22), that is the virgin, immune and suppressed state. The steady-state that one can easily predict is the suppressed state, since it has a unique value of $S/k_4 = 10$. In the suppressed state, clones have the highest population. In the virgin state clones have various populations determined by the cubic equation in x_i (3.18. The populations are the lowest and are quite similar, although only each pair of complementary clones has an identical population. In the immune state, clones can have various populations determined by the cubic equation in x_i (3.20). For this set of parameters, only one pair of complementary clones (with identical population) ended up in that state. Clones in the immune state have an intermediate population.

The state of a clone is defined by the amount of T-cell factors W_i' that it produces at equilibrium. This is shown in figure 5.13 which depicts the dynamics of the clone concentration x_i versus the T-cell factor concentration W_i'. At equilibrium, one can also see the three regions delimited by the values of $C_2 = 0.6$ and of $C_3 = 0.7$. Clones that have a low amount of T-cell factors (below 0.6) are virgin, those that have an intermediate amount of T-cell factors (between 0.6 and 0.7) are immune, and clones that have an elevated amount of T-cell factors above 0.7 are suppressed. This classification was used to label the state of the clones in figure 5.12. At equilibrium complementary clones have the same amount of T cell factors and each pair of complementary clones has its own value of the T cell factors.

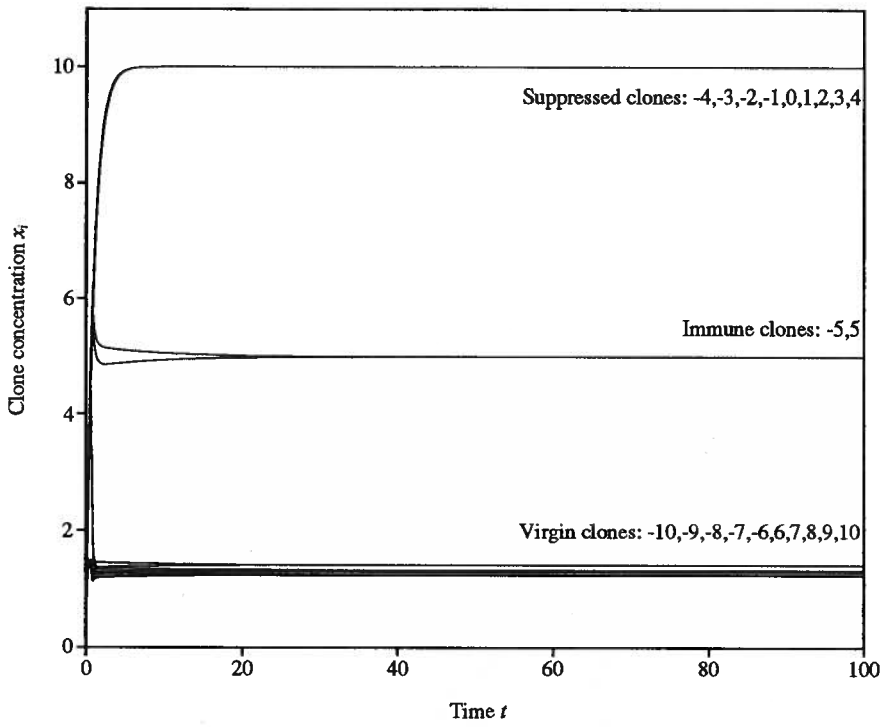

Figure 5.12: Clone concentration x_i versus time t for the non-periodic $\pm$ shape-space, with random initial conditions. $0 < x_i < 1$, $S = 10$, $k_2 = 1$, $k_3 = 0.01$, $k_4 = 1$, $C_2 = 0.6$, $C_3 = 0.7$, $\sigma = 4$, *cut-off*$= 10^{-3}$.

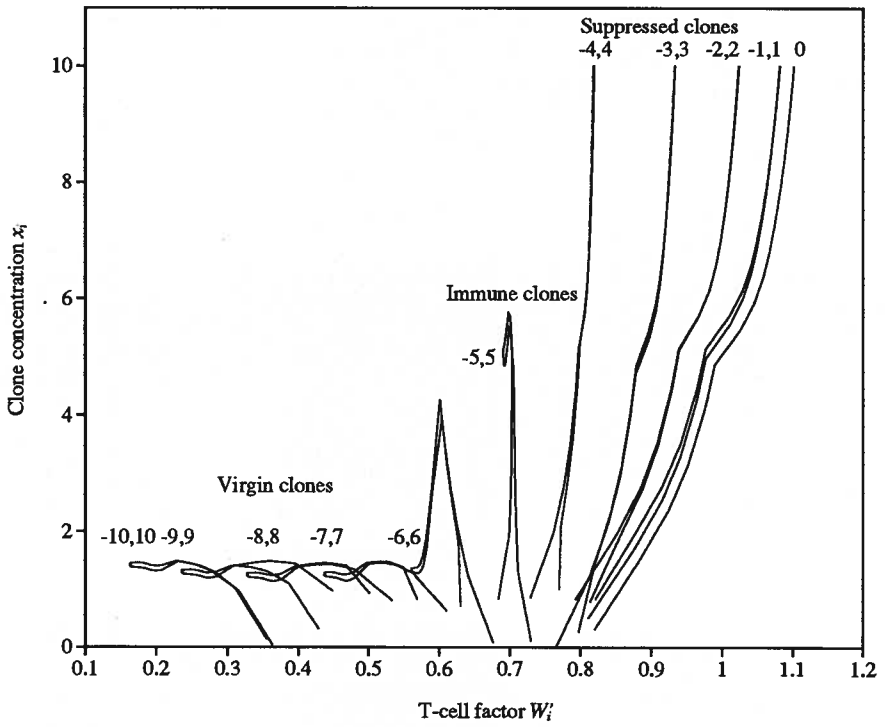

Figure 5.13: Clone concentration x_i versus T-cell factor W_i' concentration for the non-periodic $\pm$ shape-space. Same parameter values as in figure 5.12.

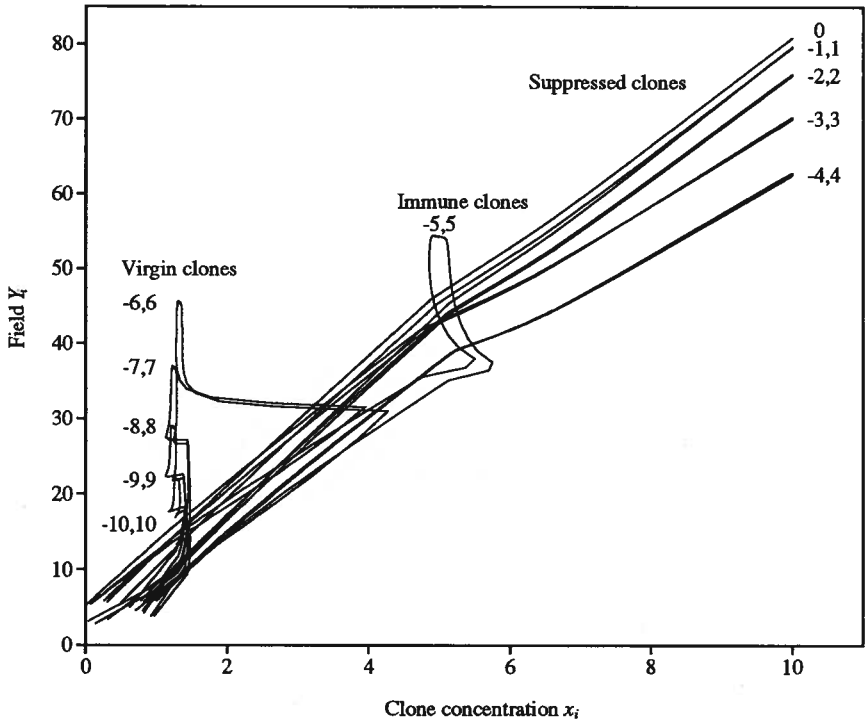

Figure 5.14: Field Y_i versus clone concentration x_i for the non-periodic $\pm$ shape-space. Same parameter values as in figure 5.12.

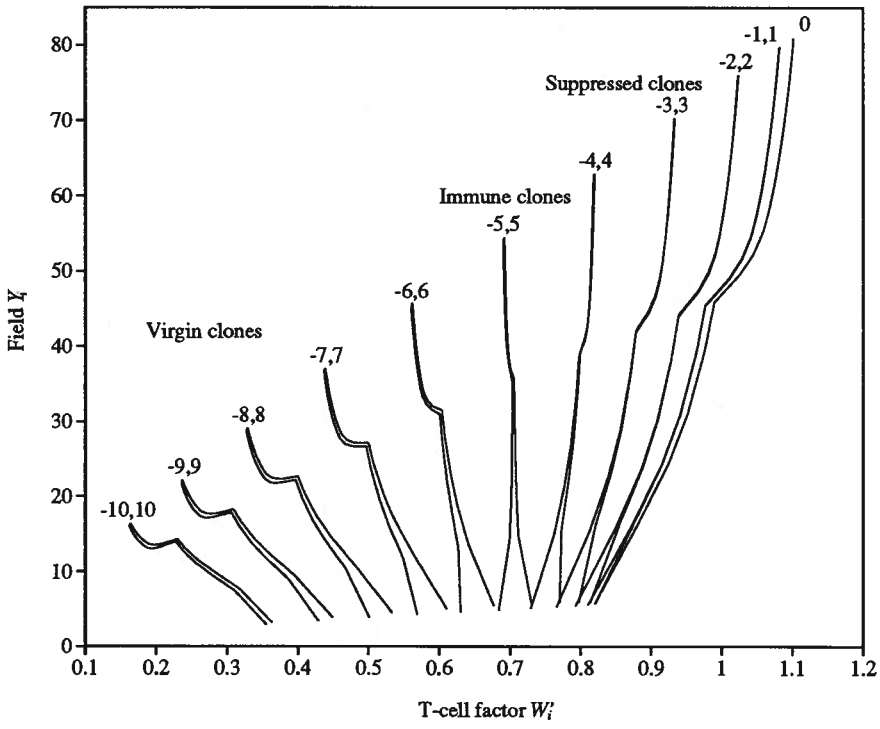

Figure 5.15: Field Y_i versus T-cell factor W_i' concentration for the non-periodic $\pm$ shape-space. Same parameter values as in figure 5.12.

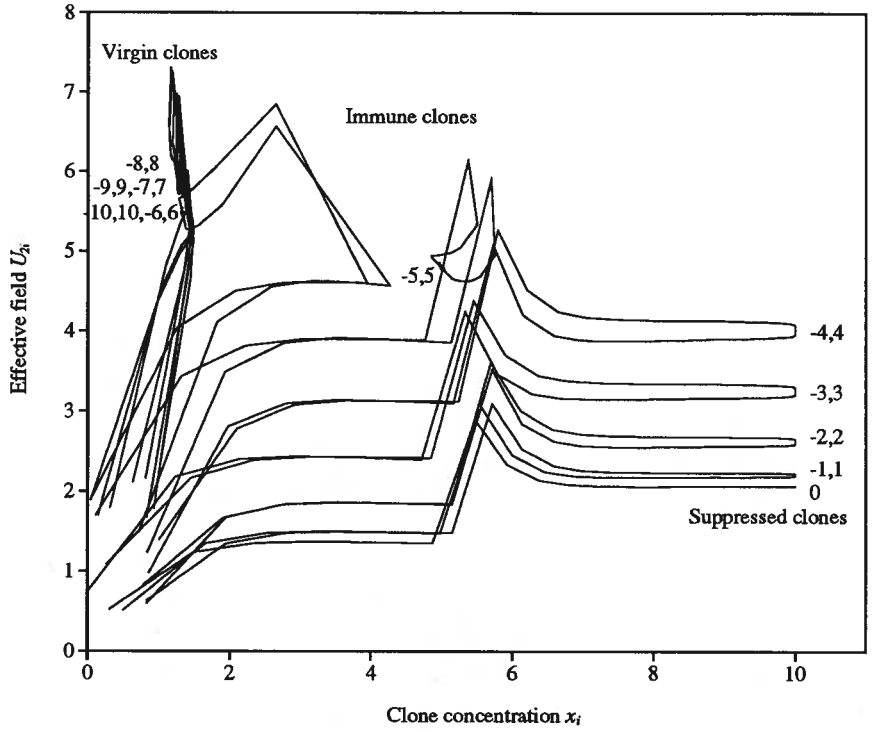

Figure 5.16: Effective field U_{2i} versus clone x_i concentration for the non-periodic $\pm$ shape-space. Same parameter values as in figure 5.12.

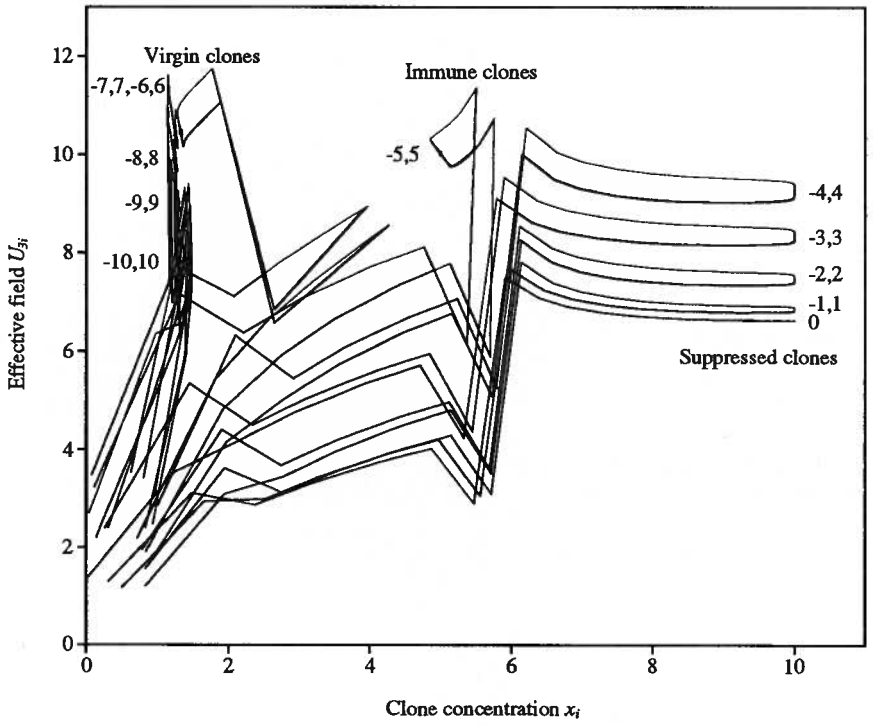

Figure 5.17: Effective field U_{3i} versus clone x_i concentration for the non-periodic $\pm$ shape-space. Same parameter values as in figure 5.12.

Figure 5.14 represents the dynamics of the field Y_i versus the clone concentration x_i. At equilibrium, one can see the three regions of values of the field. This illustrates the fact that clones with a low field are virgin , those with an intermediate field are immune, and those with a high field are suppressed. Complementary clones have the same field. Each pair of complementary clones has its own value of the field. The fields are also plotted against the T-cell factor concentrations in figure 5.15. Remember that one is somewhat dependent on the other. The graph shows a "linear" configuration of the equilibrium points.

The last variables of the model which can be of interest are the effective fields U_{2i} and U_{3i}. They are depicted in figures 5.16 and 5.17 with respect to the clone concentration. The pattern exhibited is much more dramatic, with sharp corners in the trajectories, that is with major changes in the continuity of the, at times, smooth behaviour. Complementary clones have the same effective fields. Each pair of complementary clones has its own value of effective fields.

The interest here was more to examine the configuration of the clones in shape-space, so the graphs of the dynamics will be left for now without further analysis.

Clusters in shape-space Clones are found to form clusters of the steady-states in shape-space (figure 5.18). The clones position themselves in the clusters such that clone zero is suppressed and is the center of the unique cluster of 9 suppressed clones. Then, on each side of this cluster, there is an immune clone, separating the cluster of suppressed clones from the two clusters of 5 complementary virgin clones. Complementary shapes end up in the same steady-state and have the same steady-state population. Each pair of virgin and immune complementary clones has its own steady-state population. All suppressed clones have the same steady-state population, as was shown previously in §3.5.2. It is verified analytically in appendix B.2 that the differential equation (3.4) with

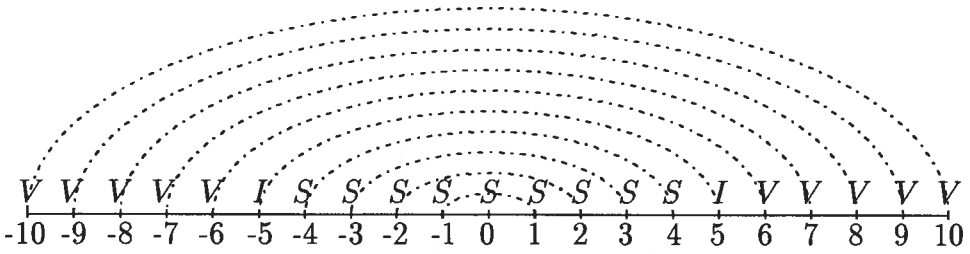

Figure 5.18: Formation of clusters of clones in the virgin(V), immune(I), suppressed(S) steady-states in the non-periodic ± shape-space for a 21-dimensional network with random initial conditions. Complementary clones share the same steady-state concentration. There is a symmetry in the configuration of clusters. Same parameter values as in figure 5.12, so $\sigma = 4$ and *cut-off*$= 10^{-3}$. This configuration is also the same for values of σ between 2 and 4.

which the affinities obey the relationship (5.25) has a solution such that complementary shapes have the same concentration.

Note that clone 0 is at the center of the cluster of suppressed clones, i.e. those that have the highest concentration. De Boer et al. [33] had observed that clone 0 autostimulates itself to grow because it has maximum complementarity with itself, and influences the other clones around it to grow. But their model involves positive stimulation terms rather than the negative suppressive terms of Hoffmann's model. So their explanation is valid in their case, but not in this one. In Hoffmann's model, the scenario is different and less obvious. The amount of T cell factors that clone 0 has initially is high enough (see figure 5.13) that both negative terms due to suppression by other clones are zero in the differential equation (3.4). Therefore, clone 0 grows until it reaches its maximum value in the suppressed state. If one looks at the matrix of affinities displayed in figure 5.11, one realizes that clone 0 is the one clone for which the truncation of the tails of the Gaussian imposed by the non-periodic boundary conditions is the least effective. As clones are positioned further away in shape-space from clone 0, the influence of the truncation of

the tails of the Gaussian increase. Clones -10 and 10 are the ones that have the smallest amount of interactions with other clones. As one can see from equations (3.6), (3.7) and (3.8), the amount of T-cell factors produced by a clone is strongly dependent on its affinities with other clones. So, even when considering the influence of the clone concentration on the computation of the T-cell factors, there should be a general tendency for the clones near clone 0 to have the greatest amount of T-cell factors. This was actually seen to be the case in the different simulations which were done for various sets of initial conditions.

Modify the random seed generator of the initial populations of clones The random seed that generates the initial clone populations was changed to different values. There was no observed change in the positions of the clusters. The values of the steady-states essentially remain the same.

Modify the variance The variance was given different values from 2 to 7. Major changes were observed. The different configurations in shape-space obtained are depicted in figures 5.18 to 5.22. For $\sigma = 2$ to 4, the configuration in shape-space was as shown in figure 5.18. It is made of clones -10 to -6 being virgin, clone -5 being immune, clone -4 to 4 being suppressed, clone 5 being immune, clones 6 to 10 being virgin. Complementary pairs of clones share the same steasy-state population in all cases. As σ is increased towards the value 4.5, there is a break in the symmetry. The concentration of the complementary clones begins to differ slightly. As it reaches 4.5, the steady-state situation becomes more asymmetric. The disparity between the concentrations of the complementary clones is greater, and clone -5 switches from being immune to being suppressed (figure 5.19). When $\sigma = 5$, the clone -6 switches from being suppressed to being immune (see figure 5.20). As σ reaches 5.5, the clone -5 returns to its original

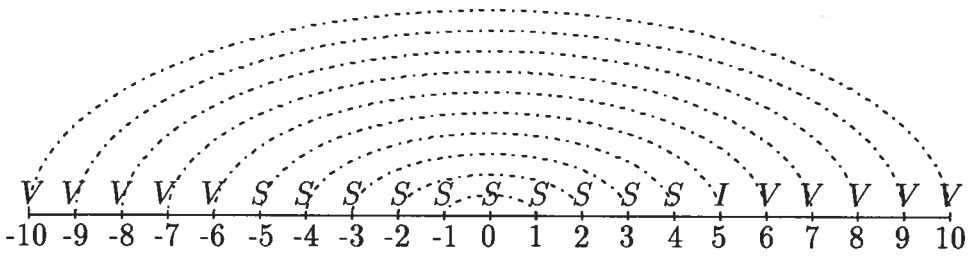

Figure 5.19: Same as in figure 5.18 but now $\sigma = 4.5$. Clone -5 has become suppressed(S). The concentrations of complementary clones are no longer identical.

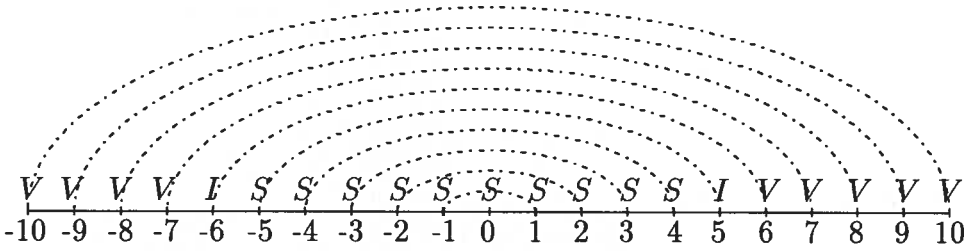

Figure 5.20: Same as in figure 5.18 but now $\sigma = 5$. Clone -6 has become immune(I). Concentrations of complementary clones are still not identical.

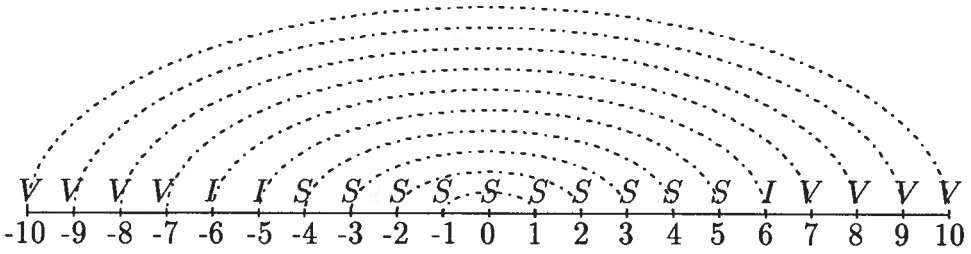

Figure 5.21: Same as in figure 5.18 but now $\sigma = 5.5$. Clone -5 and 6 have become immune(I) and clone 6 has become suppressed(S). Concentrations of complementary clones are still not identical.

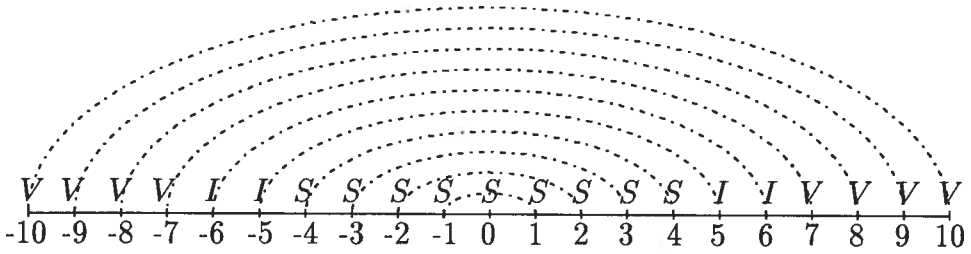

Figure 5.22: Same as in figure 5.18 but now $\sigma = 6$. Clone 5 has become immune(I), so that now clones -6, -5, 5 and 6 are immune(I). The situation is symmetric again: complementary clones have the same concentration and are in the same steady-state. The same situation holds for $\sigma = 7$.

immune state, clone 5 switches from an immune to a suppressed state and clone 6 switches from a virgin to an immune state (see figure 5.21). Then, the concentrations of the complementary clones get more similar, until at $\sigma = 6$ clone 5 returns to its original immune state (see figure 5.22). Now, the configuration of the clusters is symmetric and the concentration of complementary shapes is identical again. And we have that the immune clusters are comprised of two clones instead of one as we had at our point our departure for $\sigma = 2$. For $\sigma = 7$, the same situation holds.

The system has not been studied for greater or smaller values of the variance, so there is left an open question as to whether one could have other regions of the parameter space for which the symmetry would break and then be retrieved with a different number of immune clones.

5.2.3 The periodic ± shape-space

The system is now studied with periodic boundary conditions on the shape-space, which simulate an infinite domain. There are different ways that one could do so. One could set a period of $2N$. This was the choice of De Boer et al. [33]. This means that there

is no separation in shape-space between shape $-N$ and shape N. A period of $2N+1$ is chosen here. Figure 5.23 shows the two different periodic shape-spaces. To indicate the periodicity, "circular" diagrams of shape-spaces are now presented, that is the ends of the finite one-dimensional shape-space are joined. The matrix of affinities was computed with a Fortran program based on equation (5.25) but including several modifications to simulate the periodicity of the shape-space. It is shown in figure 5.24 for $N = 10$. Note that an implication of having periodic boundary conditions is that the sum of the affinities of one clone with the whole network is the same for every clone in the network, that is:

$$\sum_{j=1}^{N} K_{ij} = constant \qquad \text{for each clone } i. \tag{5.26}$$

The periodic ± shape-space with uniform initial conditions

The system was first studied with uniform initial conditions for the clones concentrations. It is seen that the N-dimensional problem becomes one-dimensional because of the symmetry brought about by the fact that the clones in the system encounter identical stimulation. The analytical proof for this is presented in appendix A. The main advantage of using uniform initial clone populations is that it permits one to find a set of parameters that constitutes a point of departure for our simulations. In the case of random initial clone populations, the system is too complex to allow for this finding. So the parameters for all the simulations in this work were based on the ones found with the procedure indicated in appendix A. The one-dimensional is not that interesting in itself, since all clones reach the same steady-state. We cannot see any difference in the configuration of the clusters in shape-space when we change the value of the variance, for example. We therefore restrict ourselves to non-uniform initial conditions for the rest of the numerical experiments.

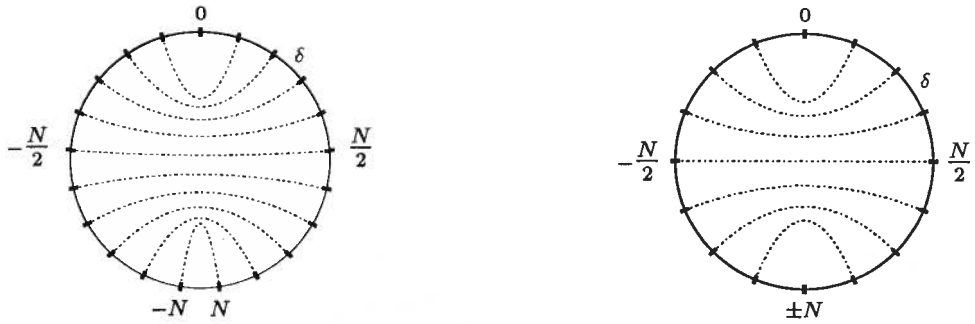

Figure 5.23: The $\pm$ shape-space with periodicity $2N+1$ (left) and with periodicity $2N$ (right). Clones are positioned on the "circular" shape-space. Each clone is most similar to its nearest neighbours. Dotted lines join the clones of maximum complementarity.

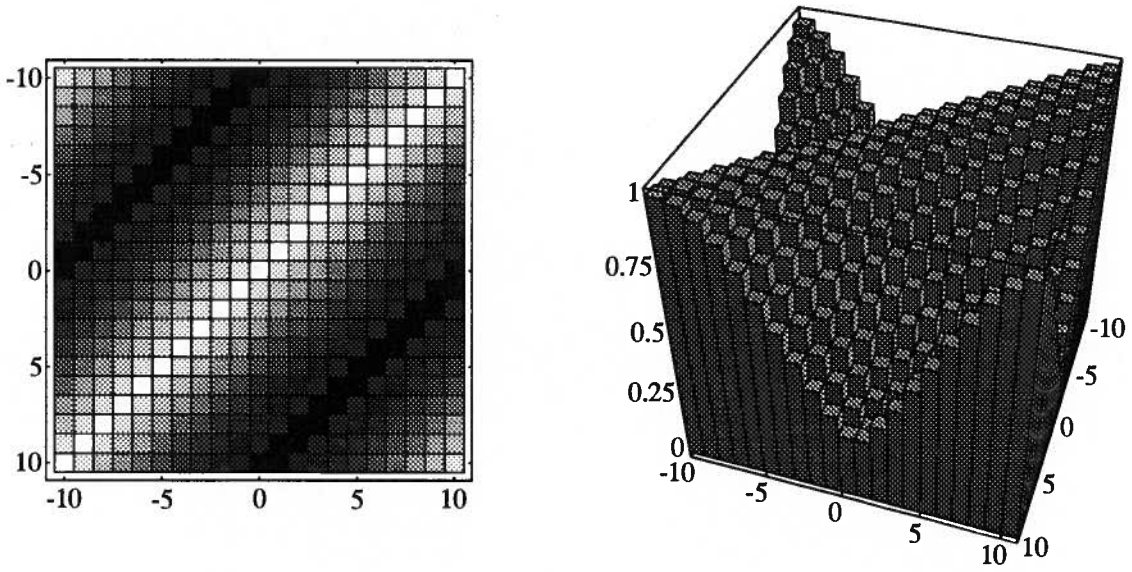

Figure 5.24: The matrix of affinities for the $\pm$ shape-space, with periodicity $2N+1$, for $\sigma = 6$ and *cut-off*$= 10^{-6}$. Refer to figure 5.11 for explanations on the legend.

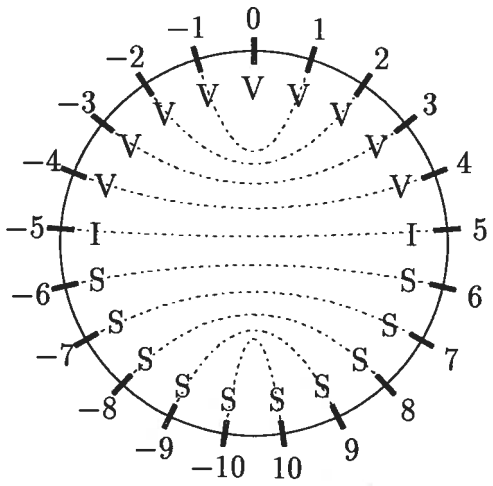

Figure 5.25: Formation of clusters in the periodic ± shape-space. This experiment was done for the exact same values of parameters as in figure 5.18, except for a different set of initial conditions. The only observed difference is the switch between clones in the virgin(V) and suppressed(S) states.

The periodic ± shape-space with random initial conditions

The system is examined for exactly the same values of parameters as in figure 5.18 of the simulations of the non-periodic case. The new steady-state configuration attained was composed uniquely of immune clones, all with the same concentration. There was a spontaneous restoration of symmetry, such that the system somehow became effectively one-dimensional. The random seed generator of the initial conditions was changed to get a situation where all three steady-states are represented and complementary clones have the same concentration. The new configuration in shape-space is shown in figure 5.25. The difference with figure 5.18 of the non-periodic ± shape-space is simply a switch in the positions of the virgin and suppressed states.

In the non-periodic case, remember that there was an asymmetry with regard to the number of interactions that the different clones of the network would have with others.

For example, clone 0 had the greatest number of interactions with others. In the periodic case, every clone has the same number of interactions with other clones. So, one would not expect the type of special configuration in the clusters that was brought by the asymmetry in the number of interactions. And indeed, for the exact same parameter values as in figure 5.18, it was seen that by changing the boundary conditions from non-periodic to periodic, one would lose that special configuration. Surprisingly, simply changing the set of initial conditions brings a special configuration again, although this time, the configuration has "reversed" itself; the clones that were previously suppressed are now virgin and vice versa.

5.3 A new shape-space without shape zero: the Δ shape-space

5.3.1 Introduction

In the $\pm$ shape-space, one shape had maximum complementarity with itself, that is the shape zero. It was thus of interest to construct a new shape-space in which there would be no shape zero and see how this would modify the configuration of the clusters of steady-states in shape-space. A new shape-space without shape zero is now presented. It will referred to as the Δ shape-space (where Δ stands for the Greek symbol "delta"), to illustrate the fact that the complementarity in it is assigned by means of a constant translation.

5.3.2 The non-periodic Δ shape-space

In the Δ shape-space, the clones are numberered from 1 to N, where N is an even number. Once again, the distance separating two clones on the shape-space determines their similarity. But this time complementarity is assigned differently. Each clone x is set maximally complementary with the clone positioned at a distance equal to $(1 + N/2)\delta$

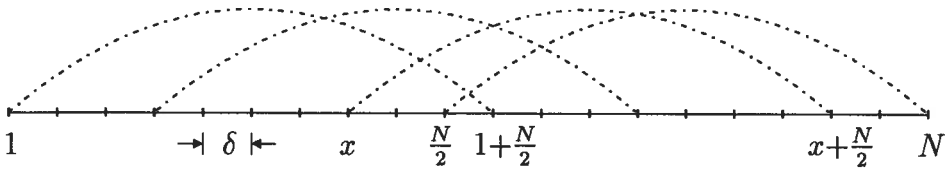

Figure 5.26: The non-periodic Δ shape-space. Clones are numbered from 1 to N (where N is an even number) and positioned on the shape-space according to the same rules of similarity with each other than for the $\pm$ shape-space. Note that the dotted lines that join clones of maximum complementarity exhibit a different pattern than for the $\pm$ shape-space.

from itself. The position of this latter clone is thus the center of the Gaussian that determines the affinities of clone x with other clones, as was the case in the $\pm$ shape-space. Figure 5.26 depicts the Δ shape-space.

Note that this particular version of shape-space is quite different from the $\pm$ shape-space depicted in figure 5.10. One can see that the dotted lines joining the clones of maximum complementarity exhibit a different pattern. As was observed in §5.2.2, in the $\pm$ shape-space there is a great flexibility about the degrees of complementarity and similarity that can be associated with two shapes. In particular, two shapes with a given distance in shape-space and that have a maximum amount of complementarity can have all degrees of similarity. But in the Δ shape-space, the rules of complementarity assign a fixed level of similarity to all pairs of maximally complementary clones.

One can also see the Δ shape-space as a reconstruction of the $\pm$ shape-space in the following way[1]. First, take the $\pm$ shape-space and remove the shape zero. Then fold the

[1]This information was given to me at the Discussion meeting for Statistical Physicists held at the St. Francis-Xavier University on October 3rd, 1993 by D. Stauffer, who himself had obtained it from S. Dasgupta [26].

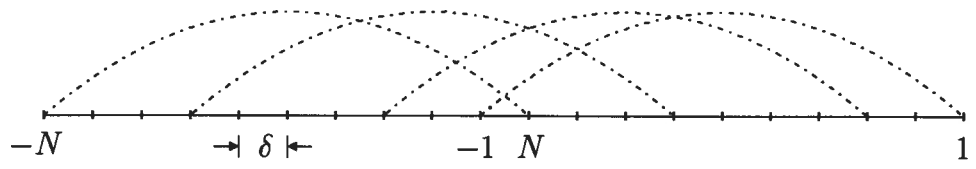

Figure 5.27: The ± shape-space with the shapes renumbered appropriately and omitting the shape zero shows the same pattern of dotted lines that join clones of maximum complementarity as in the Δ shape-space.

positive part on the negative part. That is, renumber the positive part by exchanging clone 1 with clone N, then clone 2 with clone $N-1$, etc... The reconfiguration is shown in figure 5.27. One can see that the dotted lines indicating the complementary clones are identical to the ones in the Δ shape-space.

5.3.3 The periodic Δ shape-space

We now impose periodic boundary conditions on the Δ shape-space, by joining the two ends of the shape-space together. We assign a periodicity equal to $N+1$ so that shapes 1 and N are now immediate neighbours on the shape-space. One of the consequences of doing this is that now clones of maximum complementarity are also clones of minimum similarity (or maximum dissimilarity). The periodic shape-space is shown in figure 5.28. The dotted lines that join the clones of maximum complementarity are now "diameters" of the periodic shape-space. The interactions between clones are symmetric, so the affinity between two clones i and j denoted by K_{ij} is equal to K_{ji}. To determine the $N \times N$ matrix of affinities one thus needs only specify the affinities K_{ij} between clones i and j for which $i = 1, N$ and $j = i, N$ and then symmetrize to get the rest of the matrix.

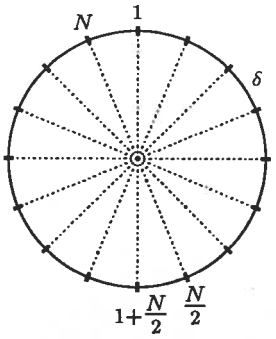

Figure 5.28: The Δ shape-space with periodicity $N+1$. Dotted lines that join the clones of maximum complementarity are diameters of the "circular" shape-space and denote that those clones are also maximally dissimilar.

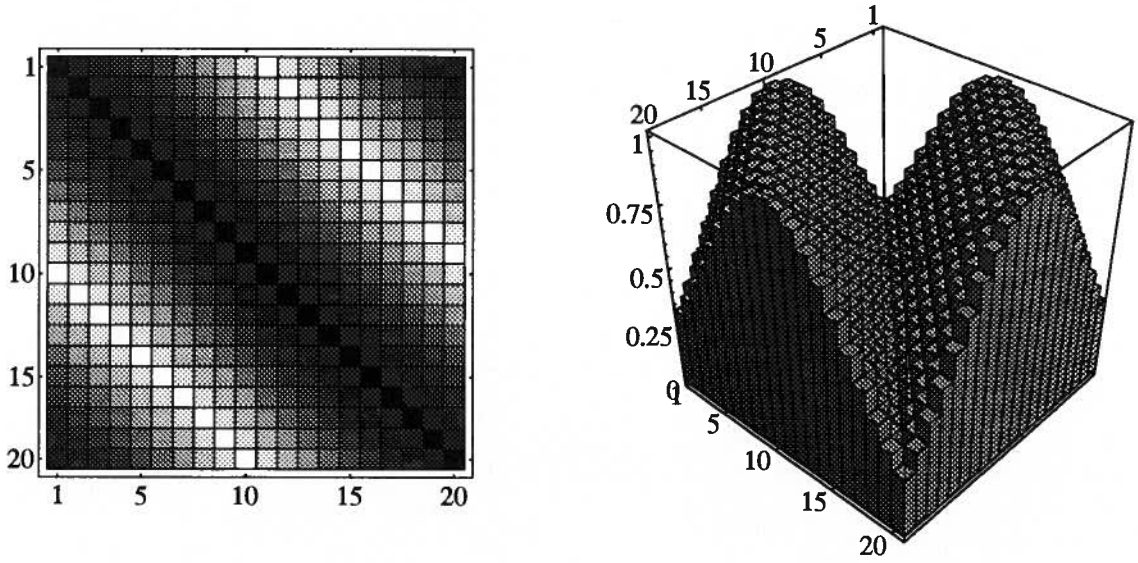

Figure 5.29: The matrix of affinities of a 20-dimensional network, for the Δ shape-space with periodic boundary conditions. The left and right pictures are two-dimensional and three-dimensional representations. Clones are numbered from 1 to 20 for both representations. In the two-dimensional representation, the shade of grey represents the value of the affinity with a scale going from white representing a maximum affinity of one and black representing a minimum affinity of zero. In the three-dimensional representation, the vertical axis represent the value of the affinity. Here, $\sigma = 6$ and cut-off= 10^{-6}.

So the affinities are defined by:

$$K_{ij} = k\ (2\pi\sigma^2)^{-1/2} exp[-(i + \frac{N}{2} - j)^2\ \delta/2\sigma^2] \tag{5.27}$$

for $i = 1, N$ and $j = i, N$, where k is a normalization constant, δ is the space between shapes on the one-dimensional shape-space axis, and σ^2 is the variance of the Gaussian. This expression is a modification on the one derived by De Boer et al. [33] for their shape-space.

The affinity matrix now has two bands of Gaussian distributions (in contrast to one band for the $\pm$ shape-space), positioned symmetrically with respect to the upper-left to lower-right diagonal, and parallel to it. It is pictured in figure 5.29 for $N = 10$.

The periodic Δ shape-space with random initial conditions

The Δ periodic shape-space has been studied only for random initial clone populations. They were randomly chosen to be between 0 and 1. The characteristic parameters of the Gaussian were $\sigma = 2$ and *cut-off*$= 10^{-3}$. A system of 20 clones is first examined.

Dynamics of the model Figures 5.30 to 5.35 show an equivalent of figures 5.12 to 5.17 of the dynamics of the different variables of the model. Figure 5.30 shows the dynamics of the clone concentrations x_i versus the time t. It can be seen that there is no great difference between figure 5.12 and figure 5.30 in that the values reached at equilibrium are essentially the same. Some difference can be observed in the way that the virgin and immune states are reached by the clones. In figure 5.12, the clone concentrations increased to a greater value than the equilibrium value, then decreased below it, then increased until they reached it and finally settled there. In figure 5.30, the clone concentration increases to a greater value than the equilibrium value, then decreases to this latter and settles there. At equilibrium, complementary clones N and

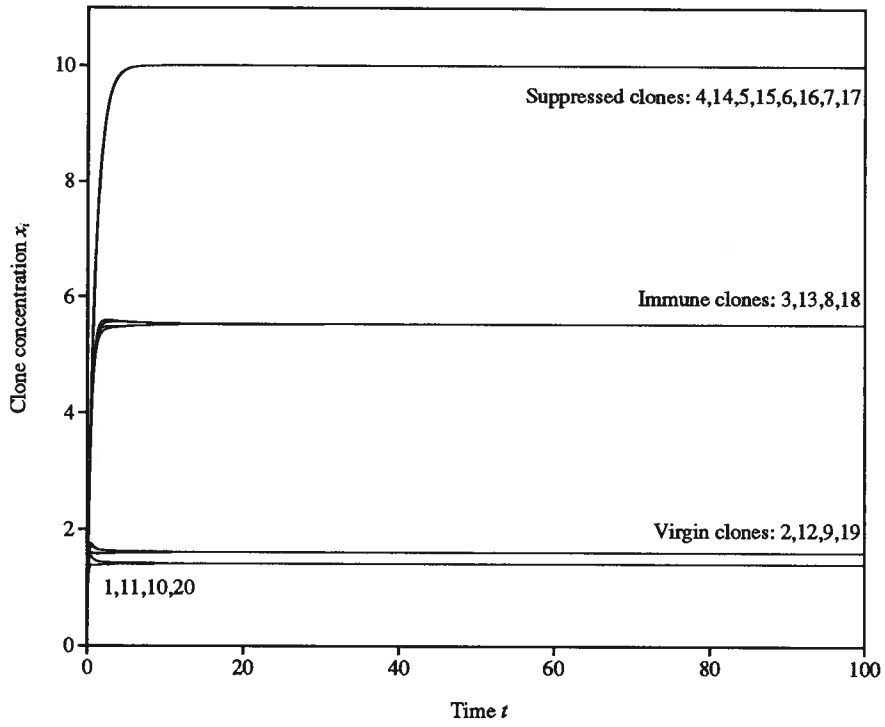

Figure 5.30: Clone concentration x_i versus time t for the periodic Δ shape-space, with random initial conditions. $0 < x_i < 1$, $S = 10$, $k_2 = 1$, $k_3 = 0.01$, $k_4 = 1$, $C_2 = 0.62$, $C_3 = 0.65$, $\sigma = 2$, *cut-off*$= 10^{-3}$.

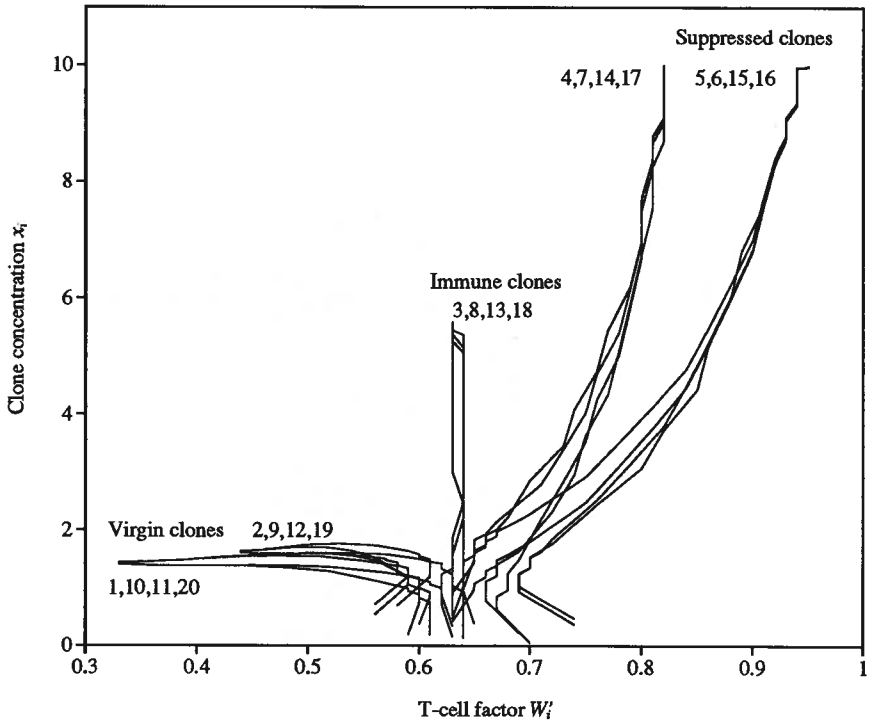

Figure 5.31: Clone x_i versus T-cell factor W_i' concentrations for the periodic Δ shape-space. Same parameter values as in figure 5.30.

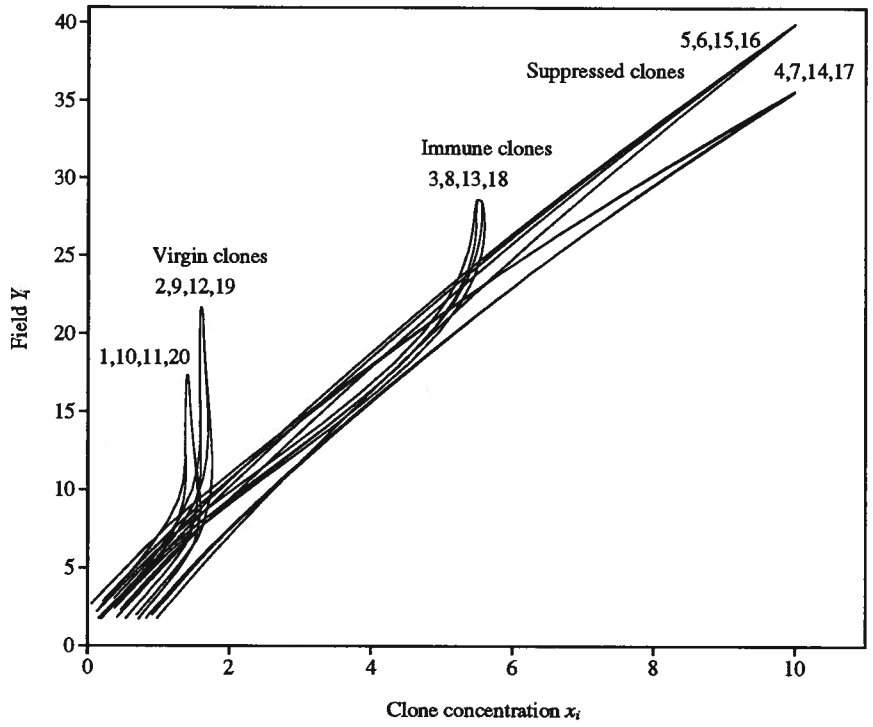

Figure 5.32: Field Y_i versus clone concentration x_i for the periodic Δ shape-space. Same parameter values as in figure 5.30.

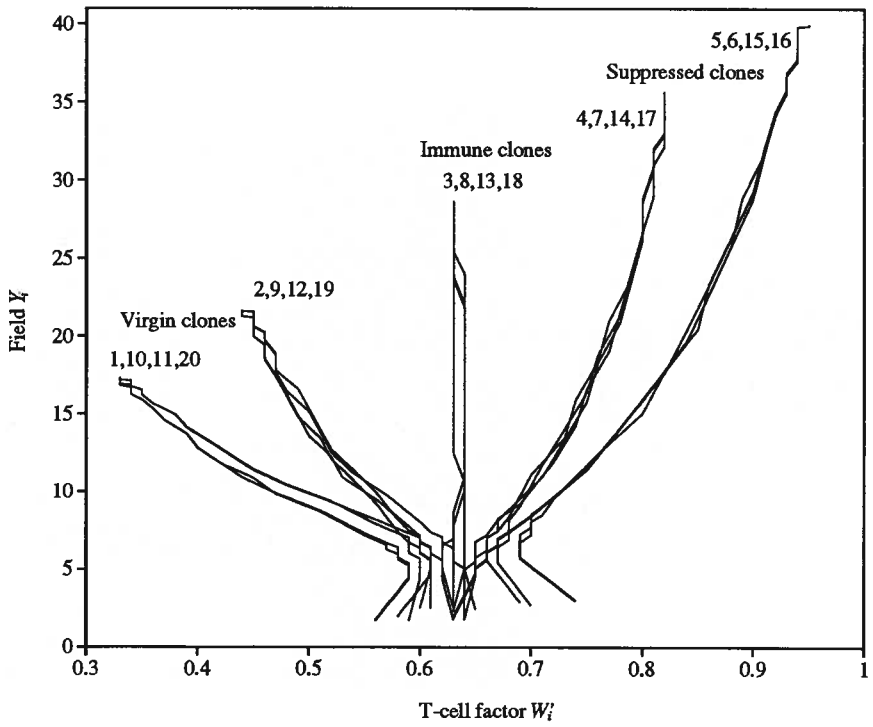

Figure 5.33: Field Y_i versus T-cell factor W_i' concentration for the periodic Δ shape-space. Same parameter values as in figure 5.30.

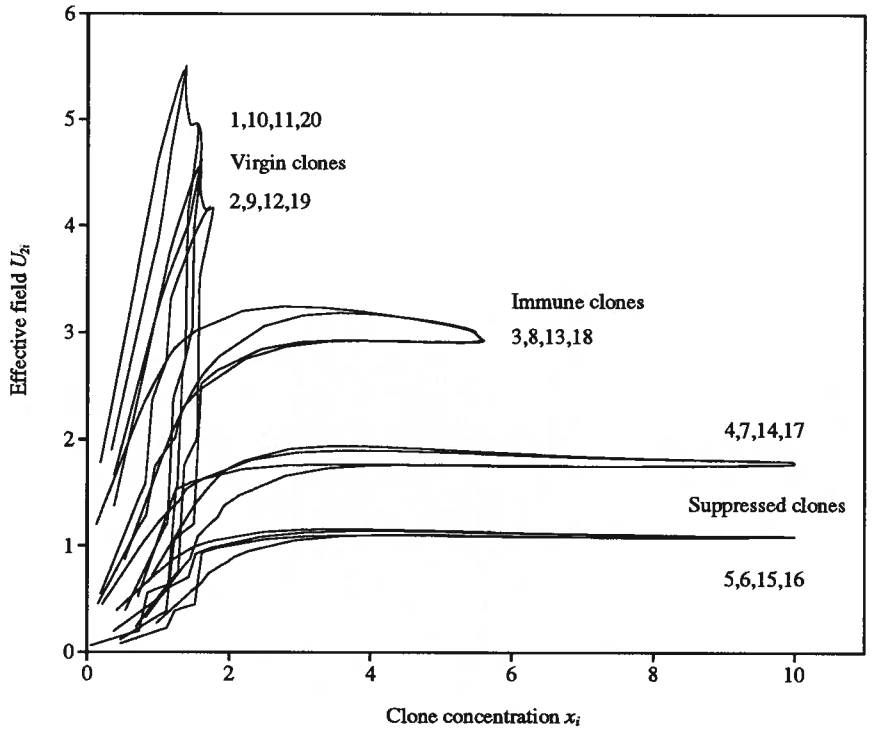

Figure 5.34: Effective field U_{2i} versus clone x_i concentration for the periodic Δ shape-space. Same parameter values as in figure 5.30.

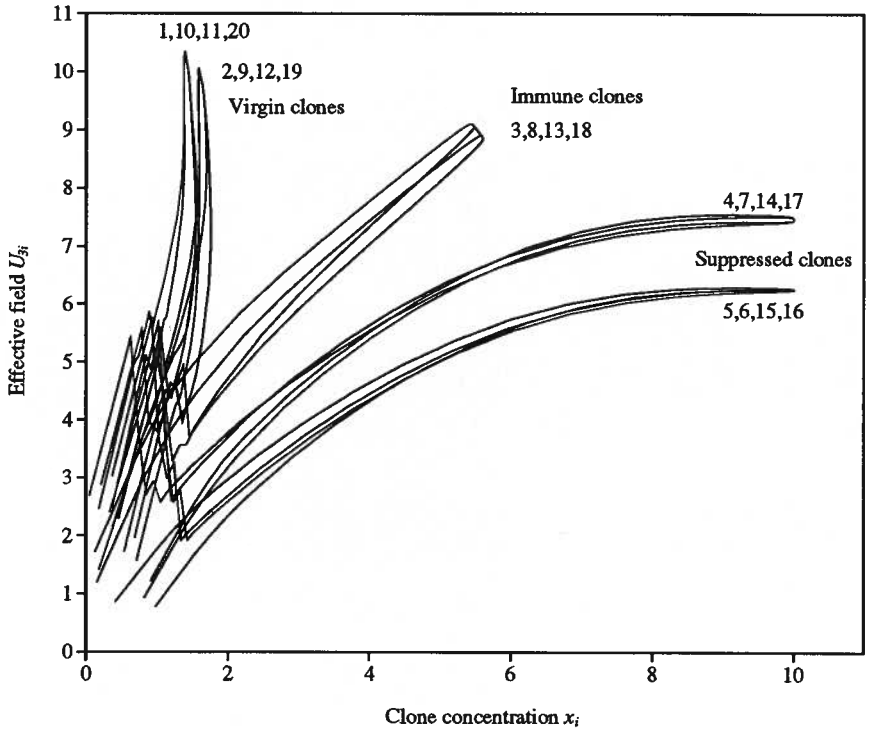

Figure 5.35: Effective field U_{3i} versus clone x_i concentration for the periodic Δ shape-space. Same parameter values as in figure 5.30.

$N + 10$ (for $N \leq 10$) have the same steady-state concentration, e.g. clones 2 and 12 are virgin with identical concentration.

Figure 5.31 depicts the dynamics of the clone concentration x_i versus the T cell factor W_i'. At equilibrium, complementary clones have the same value of T cell factors. Comparing with figure 5.13, it can be seen though that rather than each pair of complementary clones having its own equilibrium value, there are two pairs of complementary clones having the same value. Also, at time t=0, the range of values of the T cell factors is narrower in figure 5.31.

Figure 5.32 shows the dynamics of the field Y_i versus the clone concentration x_i. The same clusters of complementary clones that were sharing the same level of T cell factors (refer to figure 5.31) are seen here to share the same field. Comparing with figure 5.13, it can be seen that the behaviour is smoother here, that is not with sharp corners. Figure 5.33 shows the dynamics of the field Y_i versus the T cell factors W_i'. The greater clustering ocurring here than in figure 5.15 is again obvious. And once more it can be observed that there is a "linear" configuration of the equilibrium points.

Figures 5.34 and 5.35 depict the effective fields U_{2i} and U_{3i} versus the clone concentration x_i. At equilibrium, the same clusters of complementary clones of the previous figures are seen to have the same value of effective fields. The behaviour can also be seen to be much smoother than in figure 5.16 and 5.17. Once again, we will go directly to the study of the configuration in shape-space without examining further the graphs of the dynamical variables of the model.

The conclusions that can be drawn from comparing the non-periodic $\pm$ shape-space with the periodic Δ shape-space are: more clustering and a smoother behaviour in the latter case.

Clusters in shape-space Figure 5.38(right) shows the configuration in shape-space for a particular set of parameters. The clones end up in the three steady-states. They position themselves in a very symmetrical manner. There are two clusters of four virgin and of four suppressed clones, separated from each other by clusters of one immune clone. Complementary shapes are in the same steady-state and share the same concentration. It is shown in appendix B.1 that this is a solution.

Modifying the initial clone populations The random seed that generates the initial clone population was changed many times. The only observed change was that the configuration of clusters rotates around the shape-space. The clone populations were also given random values between 0 and 20 for the same random seed. There was no significant change. The same steady-states were occuring with essentially the same populations. The configuration of clusters stayed the same.

Modifying the values of the threshold constants Figures 5.36 to 5.39 show the configurations of the clones in shape-space as only the values of the threshold constants C_2 and C_3 are modified. Figure 5.36 left shows the configuration when $C_2 = 0.35$ and $C_3 = 0.6$. All clones are in the immune state. The concentrations of the complementary clones are not the same. For the rest of the figures, the value of C_3 is kept fixed and the value of C_2 is increased up to 0.59. In figures 5.36 (right), 5.37 (left), 5.37 (right) and 5.38 (left), the value of C_2 has been changed to 0.358, 0.36, 0.39 and 0.44. One can see that virgin clones appeared one by one consecutively at positions 8, 9, 7 and 6. They positioned themselves side by side until they formed a cluster of four clones. As C_2 reaches the value 0.45, the configuration of the shape-space changes drastically. Not only do we have one cluster of four virgin clones, but two of them, as well as two other clusters of four suppressed clones. These clusters are separated from each other by one immune

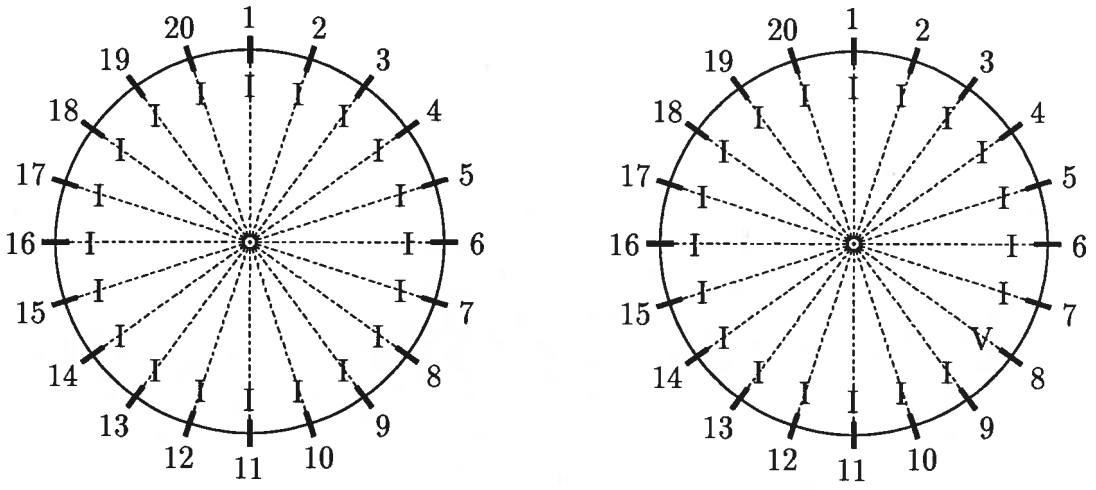

Figure 5.36: Formation of clusters of clones in the virgin(V), immune(I) and suppressed(S) steady-states in the periodic Δ shape-space for a 20-dimensional network with random initial conditions. $0 < x_i < 1$, $S = 10$, $k_2 = 1$, $k_3 = 0.01$, $k_4 = 1$, $\sigma = 2$, *cut-off*$= 10^{-3}$. Left: All clones are immune(I). Complementary clones do not share the same concentration. $C_2 = 0.35$ and $C_3 = 0.6$. Right: Clone 8 has become virgin(V). $C_2 = 0.358$ and $C_3 = 0.6$.

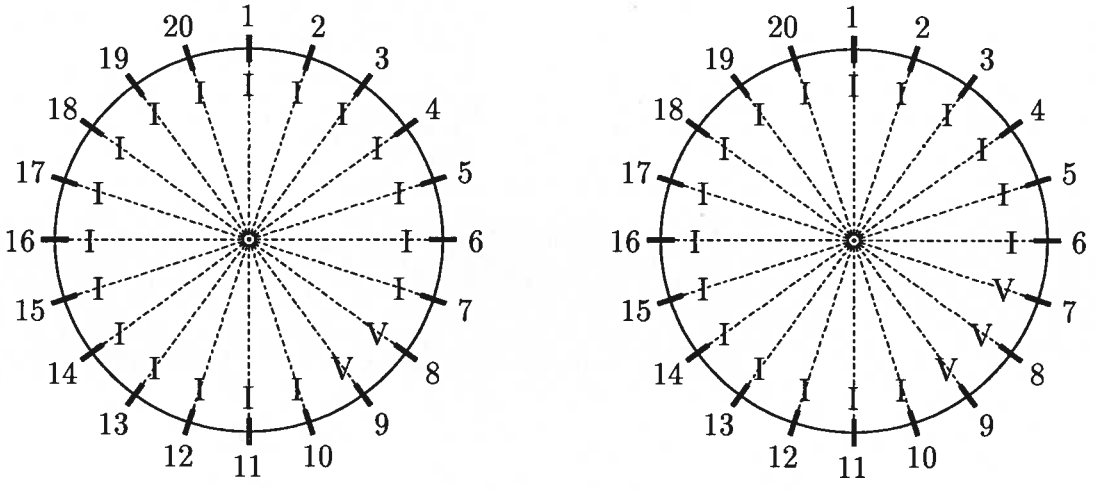

Figure 5.37: Same as in figure 5.36 excepted for the values of C_2 and of C_3. Left: Clone 9 has become virgin. $C_2 = 0.36$ and $C_3 = 0.6$. Right: Clone 7 has become virgin. $C_2 = 0.39$ and $C_3 = 0.6$.

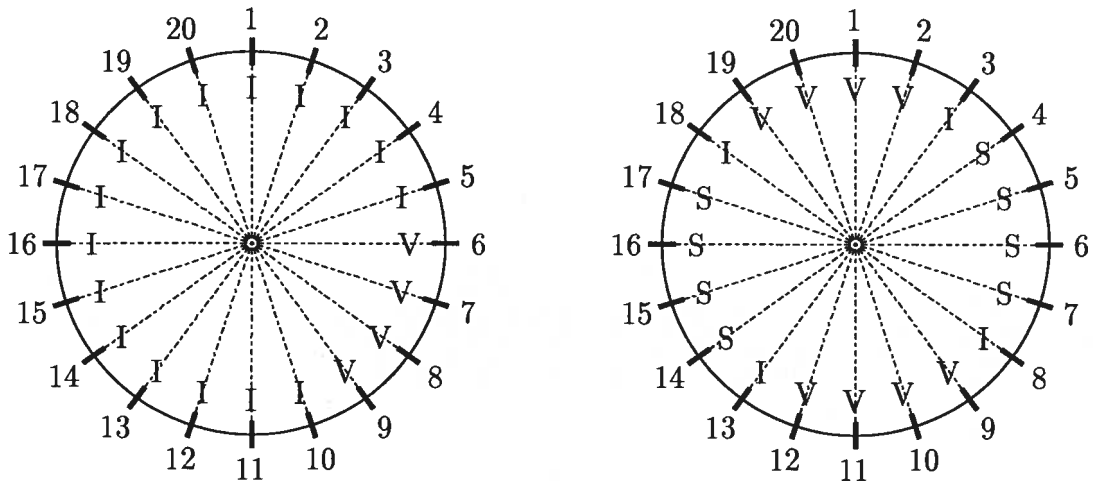

Figure 5.38: Same as in figure 5.36 excepted for the values of C_2 and of C_3. Left: Clone 6 has become virgin. $C_2 = 0.44$ and $C_3 = 0.6$. Right: Drastic change in the configuration. There are now clusters of four virgin and suppressed clones separated from each other by immune clones. The situation is symmetric and complementary clones have the same concentration. $C_2 = 0.358$ and $C_3 = 0.6$.

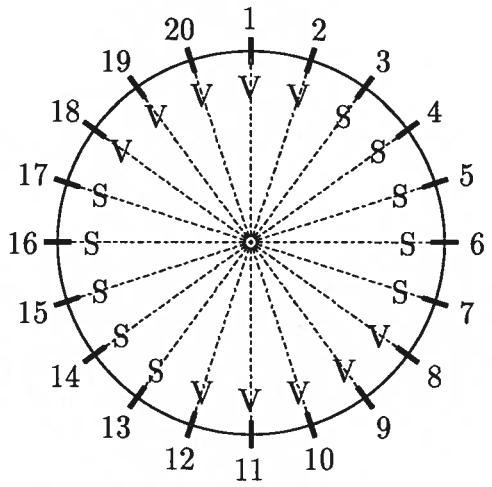

Figure 5.39: Same as in figure 5.36 excepted that $C_2 = 0.59$ and of $C_3 = 0.6$. The immune clones have disappeared and have been replaced by suppressed clones. The situation is still symmetric and complementary shapes have the same concentration.

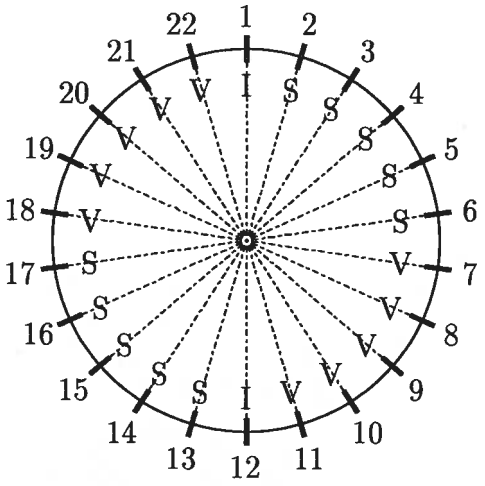

Figure 5.40: Clusters in the periodic Δ shape-space for a 22-dimensional network with random initial conditions. The configuration that exhibits all three steady-states has complementary clones sharing the same concentration and being in the same steady-state. The pattern is different than with the 20-dimensional network. Same parameter values as in figure 5.36 excepted for $C_2 = 0.65$ and $C_3 = 0.7$.

clone. The situation is symmetric and the concentrations of complementary clones are identical. It was discussed previously. Note furthermore that the original position of our first cluster of virgin clones has not been kept! As C_2 reaches the value of 0.59, the immune clones have vanished and are replaced by suppressed clones (figure 5.39). The situation is again symmetric and complementary clones share the same concentration still. But note that the fact that we have a symmetric situation does not involve having complementary shapes sharing the same concentration. In the first case all shapes were in the immune state, which can be seen as a symmetric situation, but they did not share the same concentration.

Changing the dimension of the network The nice symmetry encountered in the previous cases was thought to come from having chosen a number of 20 clones in the

network which would bring that symmetrical distribution of the clusters. To test that hypothesis, the model was studied with other numbers of clones in the network.

First, an experiment was done with 22 clones to keep the even number number of clones. The result is shown in figure 5.40. And indeed, the system exhibits a different pattern. Complementary shapes still exhibit the same concentration as it is to be expected (adding an even number of clones surely does not invalidate the demonstration presented in Appendix B.1), and there can still be three general types of steady-states. But the number of virgin and suppressed clones grouped in clusters has changed, and those clusters are not always separated from each other by an immune clone. There still are some immune clones separating some of the clusters, but one is not present and has been replaced by a suppressed clone.

The model was then studied with an odd number of clones, that is 21 clones. Each shape was set complementary to another *hypothetical* one situated between the two most complementary shapes, so that the expression giving the the affinities is:

$$K_{ij} = k\ (2\pi\sigma^2)^{-1/2} exp[(i - j + N/2 + \delta/2)^2\ \delta/2\sigma^2] \tag{5.28}$$

where $i = 1, N$ and $j = i, N$ (N is an even number), with symmetrization of the matrix. The matrix of affinities shown in figure 5.41 is very similar to the one for an even number of clones. One now has that each clone has maximum complementarity with two other clones. For example, clone 1 has maximum affinity equal to 1 with clones 11 and 12, and then a smaller affinity with clones 10 and 13, etc... The result of the simulation is shown in figure 5.42. It is found that the three actual shapes mentioned above ended up in the same virgin or suppressed state as long as the first one is part of a cluster of virgin or suppressed shapes. If the first shape ends up being in the immune state, then the two most complementary shapes are found to be one in the virgin state and the other in the suppressed state. It is believed that the hypothetical complementary shape to the actual

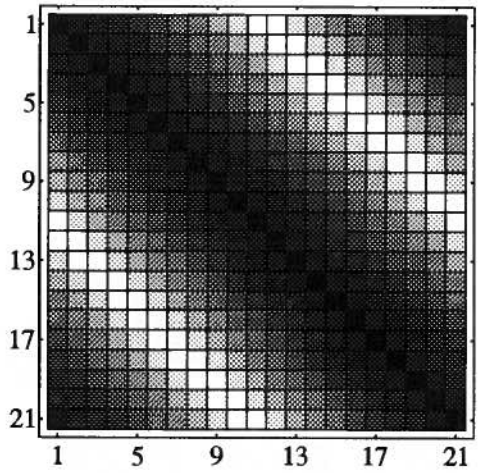

Figure 5.41: The matrix of affinities of a 21-dimensional network, for the periodic Δ shape-space. It is very similar to the one for a network comprised of an even number of clones. $\sigma = 6$ and *cut-off*= 10^{-6}.

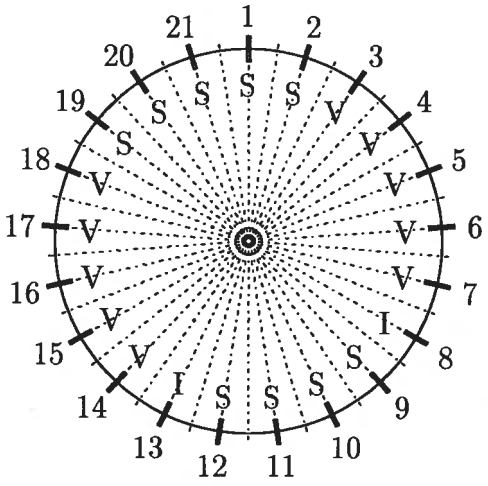

Figure 5.42: Clusters in the periodic Δ shape-space for a 21-dimensional network with random initial conditions. Dotted lines of maximum complementarity show that each clone is set complementary to a hypothetical clone situated between its most complementary clones. Same parameter values as in figure 5.36 excepted that $C_2 = 0.6$ and $C_3 = 0.7$.

shape in the immune state would also be in the immune state. The new geometry of the shape-space is such that one has actually doubled the number of actual shapes (by adding to it an equal number of hypothetical shapes), so that one ends up with a total number of an even number of clones. Then the demonstration presented in Appendix B.1 also holds for the case when one is taking into account all the actual and hypothetical shapes in it.

This ends the studies of Hoffmann's N-dimensional network in shape-space.

Chapter 6

The distance coefficient

6.1 The concept of a distance coefficient

In 1989, Hoffmann and Tufaro [62] presented the concept of a "similarity coefficient", which represents the extent to which two substances A and B are similar, as seen by a third substance C. The justification for defining such a parameter follows here. Biologists can measure experimentally the degree of complementarity of two proteins, which is defined by the affinity K. Sometimes, one wants to know how similar proteins A and B are. To answer that question, one usually compares their sequences of amino acids and DNA. But this doesn't tell much about other characteristics like the three-dimensional shapes of those proteins, which can vary significantly even for a small change in the sequence. Furthermore, different observers can have diverging opinions about the similarity of two proteins. One example given by Hoffmann and Tufaro is about the similarity of the proteins of bovine serum albumin (BSA) and of mouse serum albumin (MSA). These proteins, which have similar three-dimensional structure, would be considered as dissimilar from the point of view of a mouse or of a cow's immune system; but they would be seen as quite similar by a chicken's immune system, since chicken anti-BSA is likely to cross-react with chicken anti-MSA, the reason being the close relatedness of mice and cows from the chicken's phylogenetic point of view.

Thus it is sensible to define the similarity of two substances in the context of a third one. This can be done in various ways. One way which has been suggested by

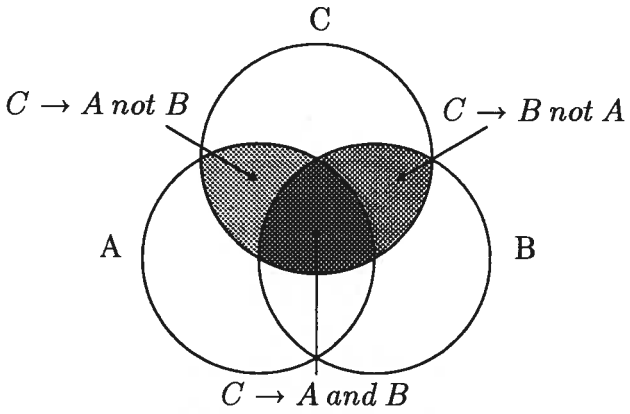

Figure 6.43: The fractions of C that are taken into account in the Hoffmann-Tufaro's definition (eq. 6.25) of the similarity coefficient. (Adapted from [62].)

Hoffmann and Tufaro is to define the similarity coefficient between two substances A and B in the context of C as the ratio between the fraction of C that reacts with A and B simultaneously and the fraction of C that reacts with either A or B or both:

$$S[\mathrm{A,B|C}] = \frac{C \rightarrow A\ and\ B}{C \rightarrow A\ not\ B\ +\ C \rightarrow B\ not\ A\ +\ C \rightarrow A\ and\ B} \tag{6.29}$$

where $C \rightarrow A\ and\ B$ denotes the fraction of C that reacts with A and B, $C \rightarrow A\ not\ B$ denotes the fraction of C that reacts with A but not with B, $C \rightarrow B\ not\ A$ denotes the fraction of C that reacts with B but not with A, and $C \rightarrow A\ and\ B$ denotes the fraction of C that reacts with both A and B. Note that this definition does not take into account the parts of C that interact with neither A or B. Figure 6.43 illustrates the different fractions of C that are taken into account in this definition. It can easily be seen that the similarity coefficient defined above is a number ranging from 0 to 1, since the denominator contains the term in the numerator plus some others. It will be zero if there is no part of C interacting with some parts of A and B simultaneously, in which case substances A and B appear completely dissimilar to substance C. It will be one if

some part of C interacts to the same extent with parts of A and B only but not with each of them separately, meaning that they are considered as being totally similar from the point of view of C. Hoffmann and Tufaro describe experimental ways of measuring this coefficient by absorption of sera against others, while being aware of the practical difficulties that can arise from it.

They also introduce the "distance coefficient" which is defined as the converse of the similarity coefficient:

$$D[\mathrm{A,B|C}] = 1 - S[\mathrm{A,B|C}] \tag{6.30}$$

$$= \frac{C \rightarrow A \, not \, B \, + \, C \rightarrow B \, not \, A}{C \rightarrow A \, not \, B \, + \, C \rightarrow B \, not \, A \, + \, C \rightarrow A \, and \, B}. \tag{6.31}$$

It also is a number ranging between 0 and 1, with properties opposite to those of the similarity coefficients'. They prove, using Venn diagrams, that the distance coefficients for sera are a metric (see Appendix C), one of the consequences being that they obey the triangle inequality such that for any three sera X, Y and Z and a reagent C, one has that:

$$D[X, Y|C] \leq D[Y, Z|C] + D[Z, X|C]. \tag{6.32}$$

The three distances between three sera can be plotted by points on a plane. They suggest it to be very useful for different practical purposes such as the diagnosis of certain diseases. Take for example the disease lupus. The diagnosis method using the distance coefficients is illustrated in figure 6.44. One measures the distance coefficient between an average serum obtained from normal individuals and an average serum of individuals for lupus, using a third reagent C. Each of the sera used to form the average sera also has a distance coefficient with each of those average sera. Ideally, when plotted in the distance coefficients plane, the normal sera will have positions somewhere close to the average normal serum and the lupus sera will be close to the average lupus serum. They

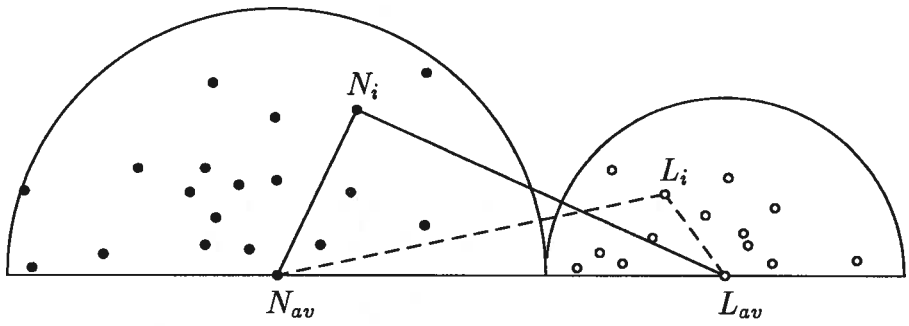

Figure 6.44: Representation in the distance coefficient plane: a useful tool for the diagnosis of diseases. N_{av} and L_{av} are the average normal and lupus sera. N_i and L_i are the individual normal and lupus sera that constitute the average sera. Plotting the distances of all these sera in the distance coefficient plane is expected to result in two different regions (illustrated with the semi-arcs). The region into which an unknown serum would be plotted could help determine whether the unknown serum is normal or from a person with lupus. (Adapted from [62].)

then define separate regions which can be used to diagnose unknown sera, depending on where the latters are located in the distance coefficients plane.

6.2 The first model of the similarity coefficient

The concept of similarity coefficient sees its first use in immune network modelling with Hoffmann et al. [61]. Choosing the context C to be the set of N idiotypes in the network, each one represented with a certain concentration, they ask how much any two idiotypes i and j taken from inside the network look similar from the point of view of the network. To answer this question, they use the original definition of the similarity coefficient presented by Hoffmann and Tufaro [62], which for this system becomes what will be called here **the first definition of the similarity coefficient**:

$$S[i,j|\mathrm{C}] = \frac{C \rightarrow i\ and\ j}{C \rightarrow i\ not\ j\ +\ C \rightarrow j\ not\ i\ +\ C \rightarrow i\ and\ j}. \tag{6.33}$$

They also present another definition which this time takes into account the fraction of C that reacts with neither clones i nor j. The basic argument is that if some clones of C react with neither clones i nor j, then this is an aspect of similarity of clones i and j. They define what will be called here **the second definition of the similarity coefficient** as the fraction of C that reacts with both i and j plus the fraction of C that reacts with neither of them, divided by the whole of C:

$$S'[i,j|\mathrm{C}] = \frac{C \rightarrow i\ and\ j\ +\ C \overset{not}{\rightarrow} i\ or\ j}{C}. \tag{6.34}$$

where $C \overset{not}{\rightarrow} i\ or\ j$ represents the fraction of C that reacts with neither A nor B, and C is the total population of C. Figure 6.45 illustrates the new fractions that are taken into account in this definition of the distance coefficients. Their work until then has been using only affinities which are dimensionless Boolean variables, i.e. if two antibodies i and j bind to each other the affinity $K_{ij} = 0$, and if they do not interact $K_{ij} = 1$. They have

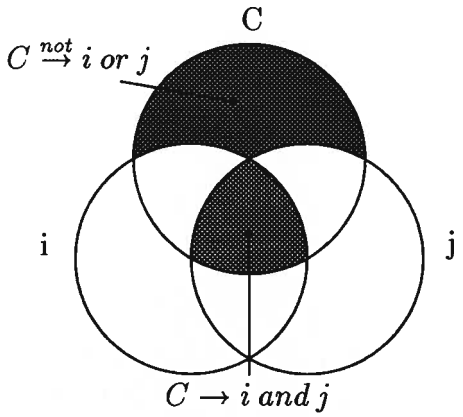

Figure 6.45: The new fractions of C that are taken into account in the second definition (eq. 6.30) of the similarity coefficient. (Adapted from [62].)

written down some equations for the similarity coefficients defined by expressions (6.33) and (6.34) using those affinities. It appears that one runs into problems when wanting to use those equations for non-Boolean affinities, i.e for $0 \leq K_{ij} \leq 1$ (where the K_{ij}'s have been normalized and non-dimensionalized). To see this, let us review the formulations of the similarity coefficient developed by Hoffmann et al. [61], using Boolean affinities.

One considers a system of N clones. One can denote the population of a clone k by x_k. One then has that:

$$C = \sum_{k=1}^{N} x_k \tag{6.35}$$

where $x_k \geq 0$. One wants to know how much the idiotypes i and j are similar from the point of view of the network. It makes sense that they will look similar to C if clones within C react to a similar extent with them. Taking a representative clone k of the system, one can write its contribution to the fraction of C that reacts simultaneously with clones i and j as the product of its affinity with each of them, weighted by its population x_k. This product is zero unless clone k reacts with both clones i and j.

Including the contributions from all the clones in the system yields the expression:

$$C \rightarrow i \; and \; j = \sum_{k=1}^{N} K_{ik} \; K_{jk} \; x_k. \tag{6.36}$$

One represents the other terms in the similarity coefficient expression in a similar manner. One can define the fraction of C that reacts with clone i but not with clone j as:

$$C \rightarrow i \; not \; j = \sum_{k=1}^{N} K_{ik} \; (1 - K_{jk}) \; x_k. \tag{6.37}$$

Each term in the summation is zero unless clone k reacts with clone i and not with clone j. Identically, one gets the fraction of C that reacts with clone j but not with clone i as:

$$C \rightarrow j \; not \; i = \sum_{k=1}^{N} K_{jk} \; (1 - K_{ik}) \; x_k. \tag{6.38}$$

The fraction of C that reacts with neither clones i nor j is expressed by:

$$C \stackrel{not}{\rightarrow} i \; or \; j = \sum_{k=1}^{N} (1 - K_{ik}) \; (1 - K_{jk}) \; x_k. \tag{6.39}$$

The term multiplying x_k in the summation is 1 only if both K_{ik} and K_{jk} are zero.

Finally, one can write the **first expression of the similarity coefficient** using (6.33) and (6.36) to (6.39), after simplifications, as:

$$S[i,j|\mathrm{C}] = \frac{\sum_{k=1}^{N} K_{ik} \; K_{jk} \; x_k}{\sum_{k=1}^{N} (K_{ik} + K_{jk} - K_{ik} \; K_{jk}) \; x_k}. \tag{6.40}$$

The **second expression of the similarity coefficient** is given by (6.34), (6.35), (6.36) and (6.39), as:

$$S'[i,j|\mathrm{C}] = \frac{\sum_{k=1}^{N} (1 - K_{ik} - K_{jk} + 2 \; K_{ik} \; K_{jk}) \; x_k}{\sum_{k=1}^{N} x_k}. \tag{6.41}$$

6.3 The limitations of the first model of the similarity coefficient

Let us here analyze these formulations. First, note that expressions (6.36) to (6.39) are correctly defined for Boolean affinities since each clone can contribute uniquely to one of the fractions of C. These fractions are all non-negative and symmetric with respect to clones i and j. Their maximum and minimum values are respectively $\sum_{k=1}^{N} x_k$ and 0. It is also easy to verify that the summation of the different fractions of C does give the whole population of C, using expressions (6.35), (6.36), (6.37), (6.38) and (6.39):

$$\begin{aligned} & C \rightarrow i \; and \; j + C \rightarrow i \; not \; j + C \rightarrow j \; not \; i + C \stackrel{not}{\rightarrow} i \; or \; j \\ &= \sum_{k=1}^{N} \left[K_{ik} K_{jk} + K_{ik} (1 - K_{jk}) + K_{jk} (1 - K_{ik}) + (1 - K_{ik})(1 - K_{jk}) \right] x_k \\ &= \sum_{k=1}^{N} x_k \\ &= C. \end{aligned} \tag{6.42}$$

Taking a close look at the first expression of the similarity coefficient given by (6.40), one realizes that it cannot be used in two cases. It will be seen that they actually do not cause any problem, since one of them appears in a situation where one is not interested in computing the similarity coefficient of two clones in the context of a network, and the other situation is non-existing for a network. If one looks at the denominator of $S[i, j|\mathrm{C}]$, one realizes that it can be zero if:

(i) all clones have a population equal to zero, that is $x_k = 0$ for all k's.

(ii) clones i and j are completely disconnected from the rest of the network, that is if $K_{ik} = K_{jk} = 0$ for all k's.

When situation (i) occurs, the network has vanished and there is no interest anymore for computation of the similarity coefficients in the context of a non-existing network.

Situation (ii) never appears since a clone considered to be part of a network must be connected to at least another clone in the network. Clones which are completely disconnected from the network do not affect the dynamics of the network and are therefore not considered to be part of it. The denominator of $S[i,j|\mathrm{C}]$ is thus always positive for all times at which one needs to compute the similarity coefficient of a network.

One can arbitrarily use the concept of a distance coefficient to demonstrate that the previous formulations of the similarity coefficient are inadequate for non-Boolean affinities. For this, only the first definition of the similarity coefficient (6.33) will be considered as a basis for deriving an expression for the distance coefficient. The conclusion derived from the following demonstration applies equally to the distance coefficient derived from the second definition of the similarity coefficient (6.34). The **first definition of the distance coefficient** derived from (6.40) and (6.30) is given by:

$$D[i,j|\mathrm{C}] = \frac{C \rightarrow i \; not \; j \; + \; C \rightarrow j \; not \; i}{C \rightarrow i \; not \; j \; + \; C \rightarrow j \; not \; i \; + \; C \rightarrow i \; and \; j}. \tag{6.43}$$

Using (6.37), (6.38) and (6.39), one can obtain the **first expression of the distance coefficient**:

$$D[i,j|\mathrm{C}] = \frac{\sum_{k=1}^{N} (K_{ik} + K_{jk} - 2\, K_{ik}\, K_{jk})\, x_k}{\sum_{k=1}^{N} (K_{ik} + K_{jk} - K_{jk}\, K_{ik})\, x_k}. \tag{6.44}$$

Since the denominator of $D[i,j|\mathrm{C}]$ is the same as the one of $S[i,j|\mathrm{C}]$, it is also positive. As pointed out by Hoffmann and Tufaro [62], the distance coefficient of a clone with itself should be zero. This makes sense biologically and is part of the requirements for the distance coefficient to form a metric. The distance coefficient of a clone i with itself

is obtained by letting $j = i$ in equation (6.44):

$$D[i,i|\mathrm{C}] = \frac{\sum_{k=1}^{N} 2\,(K_{ik} - K_{ik}\,K_{ik})\,x_k}{\sum_{k=1}^{N} (2\,K_{ik} - K_{ik}\,K_{ik})\,x_k}. \tag{6.45}$$

It can be zero *for any network* only if the affinity K_{ik} is confined to values 0 or 1. One can see a problem here. As long as one is dealing with affinities that are Boolean variables, one is fine using the formulations proposed by Hoffmann et al. [61], but as soon as one wants to use more realistic affinities whose values can range anywhere from 0 to 1, one has to find another way to express the similarity and the distance coefficients. Since only Boolean affinities have been used before in Hoffmann models, the previous formulations were suitable. The **second definition of the distance coefficient** derived from (6.41) and (6.30) is given here for completeness:

$$D'[i,j|\mathrm{C}] = \frac{C \rightarrow i\ not\ j\ +\ C \rightarrow j\ not\ i}{C}. \tag{6.46}$$

The **second expression of the distance coefficient** is thus given by:

$$D'[i,j|\mathrm{C}] = \frac{\sum_{k=1}^{N} (K_{ik} + K_{jk} - 2\,K_{ik}\,K_{jk})\,x_k}{\sum_{k=1}^{N} x_k}. \tag{6.47}$$

Note that in all of the last four formulations of the similarity and distance coefficients (6.40), (6.41), (6.44) and (6.47), the denominators are positive, for reasons mentioned above. The numerators are non-negative, since they consist of summations of the non-negative fractions of C. The coefficients are numbers ranging between 0 and 1 since their denominators consist of their numerators plus some other terms.

6.4 The first model of the distance coefficient is a metric

Note that the above demonstration has established the fact that, as long as one is using Boolean affinities, the first expression of the distance coefficient (6.43), using expressions (6.36) to (6.39) satisfy at least one of the requirements for them to be a metric, that is the second criterion stated in Appendix C. Since the only difference between the first and the second definition of the distance coefficient appears to be the denominator, it is easily seen that the second definition of the distance coefficient also satisfies the second criterion. It seems appropriate to determine whether the two definitions of the distance coefficient also satisfy the other criteria. The first of the criteria requires that the distance coefficient is never negative. This surely is true since it is a number ranging between 0 and 1. The third criterion asks for its symmetry with respect to clones i and j. This certainly is respected since the distance coefficient is composed of symmetric expressions of the fractions of C. The fourth criterion specifies that the distance coefficient be positive for distinct clones i and j. This is true since if clones i and j are distinct then there exists at least one k for which $K_{ik} \neq K_{jk}$, in which case the numerator of $D[i,j|C]$ is positive. The fifth criterion requires the distance coefficients to obey the triangle inequality. The proof of this is analogous to the proof provided by Hoffmann and Tufaro [62] for this criterion with respect to the original formulation of the distance coefficient. The original notation of Hoffmann and Tufaro will also be modified in this procedure. The total population of C can be written as:

$$c \equiv C = \sum_{k=1}^{N} x_k. \tag{6.48}$$

One can write down some expressions for the different fractions of C that react to different extents with pairs of clones i,j and j,l and i,l. The expressions that give the fractions

of C that react to different extents with clones i and j are given by:

$$\begin{aligned}
l_{ij} &\equiv C \rightarrow i\ and\ j = \sum_{k=1}^{N} K_{ik}\ K_{jk}\ x_k \\
l_i &\equiv C \rightarrow i\ not\ j = \sum_{k=1}^{N} K_{ik}\ (1 - K_{jk})\ x_k \\
l_j &\equiv C \rightarrow j\ not\ i = \sum_{k=1}^{N} K_{jk}\ (1 - K_{ik})\ x_k \\
l^*_{ij} &\equiv C \stackrel{not}{\rightarrow} i\ or\ j = \sum_{k=1}^{N} (1 - K_{ik})\ (1 - K_{jk})\ x_k
\end{aligned} \tag{6.49}$$

so that: $l_{ij} + l_i + l_j + l^*_{ij} = c$. The expressions that give the fractions of C that react to different extents with clones j and l are given by:

$$\begin{aligned}
i_{jl} &\equiv C \rightarrow j\ and\ l = \sum_{k=1}^{N} K_{jk}\ K_{lk}\ x_k \\
i_j &\equiv C \rightarrow j\ not\ l = \sum_{k=1}^{N} K_{jk}\ (1 - K_{lk})\ x_k \\
i_l &\equiv C \rightarrow l\ not\ j = \sum_{k=1}^{N} K_{lk}\ (1 - K_{jk})\ x_k \\
i^*_{jl} &\equiv C \stackrel{not}{\rightarrow} j\ or\ l = \sum_{k=1}^{N} (1 - K_{jk})\ (1 - K_{lk})\ x_k
\end{aligned} \tag{6.50}$$

so that: $i_{jl} + i_j + i_l + i^*_{jl} = c$. The expressions that give the fractions of C that react to different extents with clones i and l are given by:

$$\begin{aligned}
j_{li} &\equiv C \rightarrow l\ and\ i = \sum_{k=1}^{N} K_{lk}\ K_{ik}\ x_k \\
j_i &\equiv C \rightarrow i\ not\ l = \sum_{k=1}^{N} K_{ik}\ (1 - K_{lk})\ x_k \\
j_l &\equiv C \rightarrow l\ not\ i = \sum_{k=1}^{N} K_{lk}\ (1 - K_{ik})\ x_k \\
j^*_{li} &\equiv C \stackrel{not}{\rightarrow} l\ or\ i = \sum_{k=1}^{N} (1 - K_{lk})\ (1 - K_{ik})\ x_k
\end{aligned} \tag{6.51}$$

so that: $j_{li} + j_l + j_i + j^*_{li} = c$.

One can rewrite each of the fractions previously defined in (6.49), (6.50) and (6.51) with respect to pairs of clones i, j and j, l and i, l as sums of other fractions with respect to clones i, j, l. For this, one must first define those latter fractions:

$$\begin{aligned}
x &\equiv C \rightarrow i \; not \; j \; not \; l = \sum_{k=1}^{N} K_{ik} \, (1 - K_{jk}) \, (1 - K_{lk}) \, x_k \\
y &\equiv C \rightarrow j \; not \; l \; not \; i = \sum_{k=1}^{N} K_{jk} \, (1 - K_{ik}) \, (1 - K_{lk}) \, x_k \\
z &\equiv C \rightarrow l \; not \; i \; not \; j = \sum_{k=1}^{N} K_{lk} \, (1 - K_{ik}) \, (1 - K_{jk}) \, x_k \\
u &\equiv C \rightarrow j \; and \; l \; not \; i = \sum_{k=1}^{N} K_{jk} \, K_{lk} \, (1 - K_{ik}) \, x_k \\
v &\equiv C \rightarrow i \; and \; l \; not \; j = \sum_{k=1}^{N} K_{ik} \, K_{lk} \, (1 - K_{jk}) \, x_k \\
w &\equiv C \rightarrow i \; and \; j \; not \; l = \sum_{k=1}^{N} K_{ik} \, K_{jk} \, (1 - K_{lk}) \, x_k \\
t &\equiv C \rightarrow i \; and \; j \; and \; l = \sum_{k=1}^{N} K_{ik} \, K_{jk} \, K_{lk} \, x_k \\
s &\equiv C \stackrel{not}{\rightarrow} i \; or \; j \; or \; l = \sum_{k=1}^{N} (1 - K_{ik}) \, (1 - K_{jk}) \, (1 - K_{lk}) \, x_k.
\end{aligned} \tag{6.52}$$

It can then easily be seen that fractions (6.52) are related to fractions (6.49) to (6.51) in the following way:

$$\begin{aligned}
(w + t) + (x + v) + (y + u) + (z + s) &= l_{ij} + l_i + l_j + l^*_{ij} = c \\
(u + t) + (y + w) + (z + v) + (x + s) &= i_{jl} + i_j + i_l + i^*_{jl} = c \\
(v + t) + (x + w) + (z + u) + (y + s) &= j_{li} + j_l + j_i + j^*_{li} = c.
\end{aligned} \tag{6.53}$$

Now, consider the first expression of the distance coefficient (6.44). One can rewrite the first expressions of the distance coefficient for pairs of clones i, j and j, l and i, l that are in terms of the fractions (6.49) to (6.51 as first expressions for clones i, j, l that are in

terms of fractions (6.52):

$$\begin{aligned} D[i,j|\mathrm{C}] &= \frac{l_i + l_j}{l_i + l_j + l_{ij}} \\ &= \frac{(x+v)+(y+u)}{(x+v)+(y+u)+(w+t)} \end{aligned} \tag{6.54}$$

$$\begin{aligned} D[j,l|\mathrm{C}] &= \frac{i_j + i_l}{i_j + i_l + i_{jl}} \\ &= \frac{(y+w)+(z+v)}{(y+w)+(z+v)+(u+t)} \end{aligned} \tag{6.55}$$

$$\begin{aligned} D[i,l|\mathrm{C}] &= \frac{j_i + j_l}{j_i + j_l + j_{li}} \\ &= \frac{(x+w)+(z+u)}{(x+w)+(z+u)+(v+t)}. \end{aligned} \tag{6.56}$$

It has already been shown by Hoffmann and Tufaro (see Appendix C) that expressions (6.54), (6.55) and (6.56) satisfy the triangle inequality.

Now, consider the second expression of the distance coefficient (6.47). The only difference with the first expression (6.44) is in the denominator. Actually, the situation is even simpler now since the three expressions for the distance coefficients $D[i,j|\mathrm{C}]$, $D[j,l|\mathrm{C}]$ and $D[i,l|\mathrm{C}]$ now have the same denominator, that is the whole of C. Then, it suffices to demonstrate that the sum of the numerators for two clones is greater than the numerator of one clone. So one one wants to prove that:

$$\begin{aligned} (l_i + l_j) + (i_j + i_l) &\geq (j_i + j_l) \\ [(x+v)+(y+u)] + [(y+w)+(z+v)] &\geq [(x+w)+(z+u)]. \end{aligned}$$

But this last equation is equivalent to say that $2\,(v+y) \geq 0$, which is certainly true, since $v, y \geq 0$. Thus, it is seen here that the second expression of the distance coefficient is also satisfying the requirements for it to satisfy the triangle inequality.

6.5 The generalized distance coefficient

The following is a proposition for the generalization of the old expressions (6.40), (6.41), (6.44) and (6.47) towards their utilization with non-Boolean variables. It will be seen that new expressions can be found for each of the fractions of C expressed in (6.36) to (6.39). Those expressions will not be truly describing these fractions. Remember that in the Boolean case, a clone in the system contributes uniquely to one of the fractions of C. With non-Boolean affinities, the problem has to be approached differently. There are of course several ways that this can be done. One simple way that one could think of to be able to use the old formulations with non-Boolean affinities is to set all affinities smaller than a certain threshold equal to an affinity of zero and all affinities greater or equal to that threshold equal to an affinity of one. But if one wants some finer "tuning", then one has to come up with another expression which achieves it. The one presented here was chosen because it was the simplest that came to mind that could be an extension of the Boolean case and for which the general behaviour was concordant with expectations. In this formulation, each clone will contribute partly to each of the different new "fractions" of C, depending on its degree of binding with the clones of C. Although it is understood that one does not deal with real fractions anymore, they will still be denoted as such, since they turn out to be real fractions in the Boolean case.

A model is considered in which the affinity K_{ij} of the clone i V region (variable region) for the clone j V region is a variable that can take any value in the range 0 to 1, that is: $0 \leq K_{ij} \leq 1$. Now one has more than two different degrees to which antibodies can bind to each other. As before, two antibodies can bind to a third one with the same degree and thus appear similar to it. For example, if $K_{ik} = K_{jk} = 0.3$, one can say that clones i and j bind to clone k with the same degree and that they should appear similar to the latter. But if one has the case where $K_{ik} = 0.3$ and $K_{jk} = 0.8$, then from the point

of view of clone k, clones i and j will appear to a certain point similar or dissimilar. The extreme case is the Boolean case, when $K_{ik} = 0$ and $K_{jk} = 1$ (or the converse). Clone k would then see clones i and j as maximally dissimilar. Thus, the contribution to $C \rightarrow i \; and \; j$ is expected to depend not only on the degree of binding of each clone i or j to clone k (see later), but also on the difference between those. If the affinities K_{ik} and K_{jk} are "close" to each other, in other words if the difference between them is small, then one would expect to have a bigger contribution to the similarity of clones i and j then if they were "far" apart.

The following expression is proposed to represent the extent to which clones i and j similarly bind to clone k:

$$similar \; extent = 1 - \mid K_{ik} - K_{jk} \mid . \tag{6.57}$$

Notice that if the difference between the affinities is small, i.e. if clones i and j bind similarly to clone k, then the extent is closer to 1. But if the difference is big then the extent is closer to 0. The similar extent is a function whose isoclines in the K_{jk} versus K_{ik} plane are given by: $K_{jk} = K_{ik} + s$, where $-1 \leq s \leq 1$. Its maximum and minimum values are respectively 1 and 0. One here would be tempted to define the contribution of clone k to $C \rightarrow i \; and \; j$ as simply equivalent to the similar extent. And indeed it could make sense that all the clones that react similarly to clones i and j should have the same contribution. But consider a specific example, with clone m which has affinities $K_{im} = 0.1$ and $K_{jm} = 0.4$, and clone n which has affinities $K_{in} = 0.6$ and $K_{jn} = 0.9$. The similar extent is the same for both clones m and n. Nevertheless, it seems reasonable that clone n should have a greater contribution, since it binds more strongly with clones i and j than does clone m. To account for that fact, one can multiply the similar extent by the product of the affinities $K_{ik}K_{jk}$. The final expression to describe the fraction of

C that reacts "similarly" to clones i and j is then given by combining (6.36) with (6.57):

$$C \rightarrow i \; and \; j = \sum_{k=1}^{N} K_{ik} \; K_{jk} \; (1 - \mid K_{ik} - K_{jk} \mid) \; x_k. \tag{6.58}$$

This surely is not a unique way that one could rewrite $C \rightarrow i \; and \; j$. No biological basis indicates that one should have a cubic expression to describe it. But as was mentioned previously, this expression is essentially the simplest that is consistent with the old Boolean expression (6.36).

In a similar manner, we can find new ways to express $C \rightarrow i \; not \; j$ and $C \rightarrow j \; not \; i$. Consider the latter. One expects the converse to the $C \rightarrow i \; and \; j$ case to be applicable in terms of similarity of clones i and j with respect to clone k. For example, if K_{ik} is close to K_{jk}, one wants the contribution to $C \rightarrow j \; not \; i$ to be small. This gives the following expression for the extent to which clones i and j bind dissimilarly with clone k:

$$dissimilar \; extent = \mid K_{ik} - K_{jk} \mid . \tag{6.59}$$

The dissimilar extent has isoclines in the K_{jk} versus K_{ik} plane that look the same as the ones of the similar extent. Its maximum and minimum values respectively are 1 and 0. Now, consider clone m with affinities $K_{im} = 0.2$ and $K_{jm} = 0.9$ and clone n with affinities $K_{in} = 0.9$ and $K_{jn} = 0.2$. The dissimilar extent is the same for both clones m and n. However, clone m should have a greater contribution to $C \rightarrow j \; not \; i$ than clone n, firstly because it binds more strongly with clone j and secondly because it binds less strongly with clone i. To represent this, one can multiply the dissimilar extent of a clone k with the product $K_{jk}(1 - K_{ik})$. Combining (6.38) and (6.59), one can then write:

$$C \rightarrow j \; not \; i = \sum_{k=1}^{N} K_{jk} \; (1 - K_{ik}) \mid K_{ik} - K_{jk} \mid x_k. \tag{6.60}$$

Identically, using (6.37) with (6.59) gives:

$$C \rightarrow i \; not \; j = \sum_{k=1}^{N} K_{ik} \; (1 - K_{jk}) \mid K_{ik} - K_{jk} \mid x_k. \tag{6.61}$$

To find the fraction of C that reacts with neither i nor j is not as intuitively simple. But one can use the argument that all the fractions of C have to add up to the whole population of C to calculate it. This yields:

$$C \overset{not}{\rightarrow} i \; or \; j = \sum_{k=1}^{N} [1 - K_{ik} \; K_{jk} - (K_{ik} + K_{jk} - 3 \; K_{ik} \; K_{jk}) \; |K_{ik} - K_{jk}|] \; x_k. \qquad (6.62)$$

The term multiplying x_k in the summation takes the value 1 if both K_{ij} and K_{jk} are 0, and the value 0 if they are both 1 or if one is 0 and the other is 1.

An a priori conceivable alternative is that that the different fractions could be written more simply in terms of the similar and dissimilar extents. In this view, one would have defined $C \rightarrow i \; and \; j$ as identical with the similar extent, and each of $C \rightarrow i \; not \; j$ and $C \rightarrow j \; not \; i$ as identical with the dissimilar extent. $C \overset{not}{\rightarrow} i \; or \; j$ would however then have to take the negative value of $-|K_{ik} - K_{jk}|$, which is not acceptable. As defined above, expressions (6.58) and (6.60) to (6.62) are all non-negative and symmetric with respect to clones i and j. Their maximum and minimum values respectively are $\sum_{k=1}^{N} x_k$ and 0.

The **generalized first expression of the similarity coefficient** can then be written, after a few simplifications, as:

$$S[i,j|\mathrm{C}] = \frac{\sum_{k=1}^{N} K_{ik} \; K_{jk} \; (1 - \mid K_{ik} - K_{jk} \mid) \; x_k}{\sum_{k=1}^{N} [K_{ik} \; K_{jk} + (K_{ik} + K_{jk} - 3 \; K_{ik} \; K_{jk}) \mid K_{ik} - K_{jk} \mid] \; x_k}. \qquad (6.63)$$

Its counterpart defining the distance coefficient gives the **generalized first expression of the distance coefficient**

$$D[i,j|\mathrm{C}] = \frac{\sum_{k=1}^{N} (K_{ik} + K_{jk} - 2 \; K_{ik} \; K_{jk}) \mid K_{ik} - K_{jk} \mid x_k}{\sum_{k=1}^{N} [K_{ik} \; K_{jk} + (K_{ik} + K_{jk} - 3 \; K_{ik} \; K_{jk}) \mid K_{ik} - K_{jk} \mid] \; x_k}. \qquad (6.64)$$

Similarly, the **generalized second expression of the similarity coefficient** is given

by:

$$S'[i,j|\mathrm{C}] = \frac{\sum_{k=1}^{N} [1 + (2\ K_{ik}\ K_{jk} - K_{ik} - K_{jk})\ |K_{ik} - K_{jk}|]\ x_k}{\sum_{k=1}^{N} x_k}. \tag{6.65}$$

And the **generalized second expression of the distance coefficient** is:

$$D'[i,j|\mathrm{C}] = \frac{\sum_{k=1}^{N} [(K_{ik} + K_{jk} - 2\ K_{ik}\ K_{jk})\ |K_{ik} - K_{jk}|]\ x_k}{\sum_{k=1}^{N} x_k}. \tag{6.66}$$

Note that all the above new expressions are generalizations of the old expressions in the sense that they reduce to the old expressions when one uses Boolean affinities.

Since expressions (6.63) to (6.66) are derived from expressions (6.58) and (6.60) to (6.62), they are numbers ranging between 0 and 1. Therefore, the distance coefficient is never negative. The dissimilar extent of a clone with itself being zero, so is the distance coefficient of a clone with itself. Since the expressions of the fractions of C are symmetric with respect to clones i and j, the distance coefficient derived from them is symmetric too with respect to those clones. For two clones i and j to be distinct, there must be at least one k for which $K_{ik} \neq K_{jk}$. Within these conditions, $D[i,j|C]$ is always positive. The first four criteria of the metric are thus fulfilled. As for the fifth criterion which is the triangle inequality, it still needs be verified analytically. But it was found numerically that for all the studied cases, the generalized distance coefficients did satisfy the triangle inequality, that is, one could always plot the distance coefficients of all clones in the network (which was the primary goal).

6.6 The distance coefficients plane

In this section, it will be demonstrated how one can plot the positions of the clones in a distance coefficient plane, once one has calculated the distance coefficients of all clones

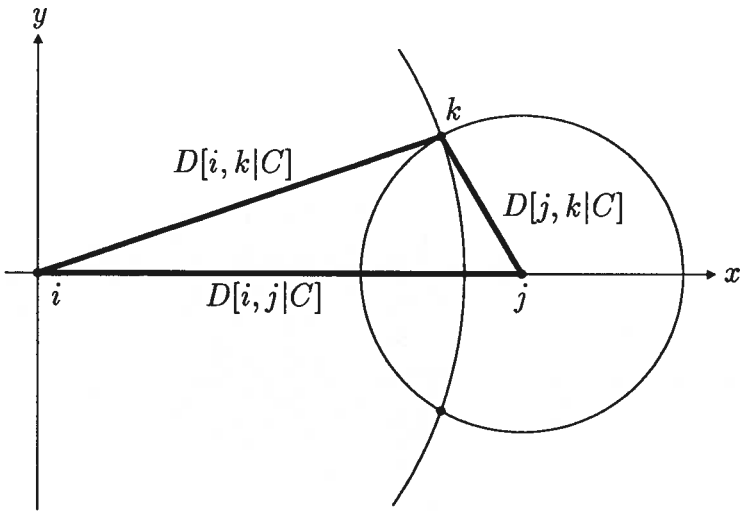

Figure 6.46: The distance coefficients plane

with respect to two arbitrary reference clones.

In the XY plane, choose the X-axis as the reference axis. Pick the two reference clones i and j. Position clone i at the origin and clone j at a distance $D[i,j|C]$ from clone i on the X-axis. One has thereby set the reference distance. Now one wants to determine the position of clone k in the plane, with respect to the reference clones. Clone k could be positioned anywhere on a circle of radius $D[i,k|C]$ centered on clone i at the origin. The equation describing the circle is given by:

$$x^2 + y^2 = D[i,k|C]^2. \tag{6.67}$$

Clone k also has to be positioned anywhere on a circle of radius $D[k,j|C]$ centered on clone j, described by:

$$(x - D[i,j|C])^2 + y^2 = D[j,k|C]^2. \tag{6.68}$$

Solving for both equations simultaneously gives the position of clone k in the plane:

$$x = \frac{D[i,k|C]^2 - D[k,j|C]^2 + D[i,j|C]^2}{2\ D[i,j|C]} \tag{6.69}$$

$$y = (D[i,k|C]^2 - x^2)^{1/2}. \tag{6.70}$$

where the positive square root is chosen. Since the two circles intersect symmetrically at two different places above and below the X-axis, one can choose either.

6.7 Simulations of the distance coefficient

6.7.1 Affinities from the periodic Δ shape-space

To test the generalized formulations of the distance coefficient, one can begin by choosing a particular set of affinities relating the clones in the network. Affinities derived from the Δ shape-space with periodic boundary conditions were chosen. A network of 20 clones was studied.

Uniform initial conditions The system with uniform initial conditions is one-dimensional as described in appendix A. The distance coefficients can be plotted independently of the populations of the clones in this system. Consider the expression (6.64) giving the distance coefficients. For the uniform initial conditions prevalent in the one-dimensional system, the concentrations of all the clones are identical, and can be factorized in front of the summations. They then cancel each other in the numerator and in the denominator, and one gets the following expression for the distance coefficient:

$$D[i,j|C] = \frac{\sum_{k=1}^{N} (K_{ik} + K_{jk} - 2\ K_{ik}\ K_{jk}) \mid K_{ik} - K_{jk} \mid}{\sum_{k=1}^{N} [K_{ik}\ K_{jk} + (K_{ik} + K_{jk} - 3\ K_{ik}\ K_{jk}) \mid K_{ik} - K_{jk} \mid]}. \tag{6.71}$$

Thus in this particular system, the distance coefficient is constant in time and does not depend on the change of the populations of the clones. For a shape-space with a high level of symmetry, the individual positions of the clones in a distance coefficients plot would be expected to have a corresponding degree of symmetry. If one chooses two complementary shapes as reference shapes, one would expect that all the other clones are positioned in a symmetrical way with respect to them. For example, here clones 1 and 11 are chosen as reference clones. Then clones 6 and 16 are "perpendicular" to them in the circular shape-space (see figure 5.39); they are the most dissimilar from clones 1 and 11 for this shape-space. They are thus most "distant" from clones 1 and 11 in the distance coefficient representation. This is shown in figure 6.47, which shows the distance coefficient representation of all 20 clones with complementary clones 1 and 11 as reference shapes. The two dashed curves are part of two circles of radius 1 and of centers equal to the positions of the reference shapes. Thus, the distance coefficients have a maximum value of 1, as was shown in §6.5.

Random initial conditions The system was studied with random initial conditions, $0 < x_i < 1$. A fixed set of initial conditions and a fixed value of the variance was given, but the value of the *cut-off* was varied. The values of the threshold constants C_2 and C_3 were chosen so that in all cases, the three steady-states are present and the complementary shapes have the same concentrations. The other values of the parameters are the same as for all previous simulations in shape-space: $S = 10$, $k_2 = 1$, $k_3 = 0.01$, $k_4 = 1$.

Figures 6.48 to 6.51 show the distance coefficient dynamics for the same variance $\sigma = 3$ and different values of the *cut-off* and of the threshold constants C_2 and C_3. The initial positions of the clones are represented with the "stars". Their trajectories in time are along the full lines. Their final positions at equilibrium is depicted by one of

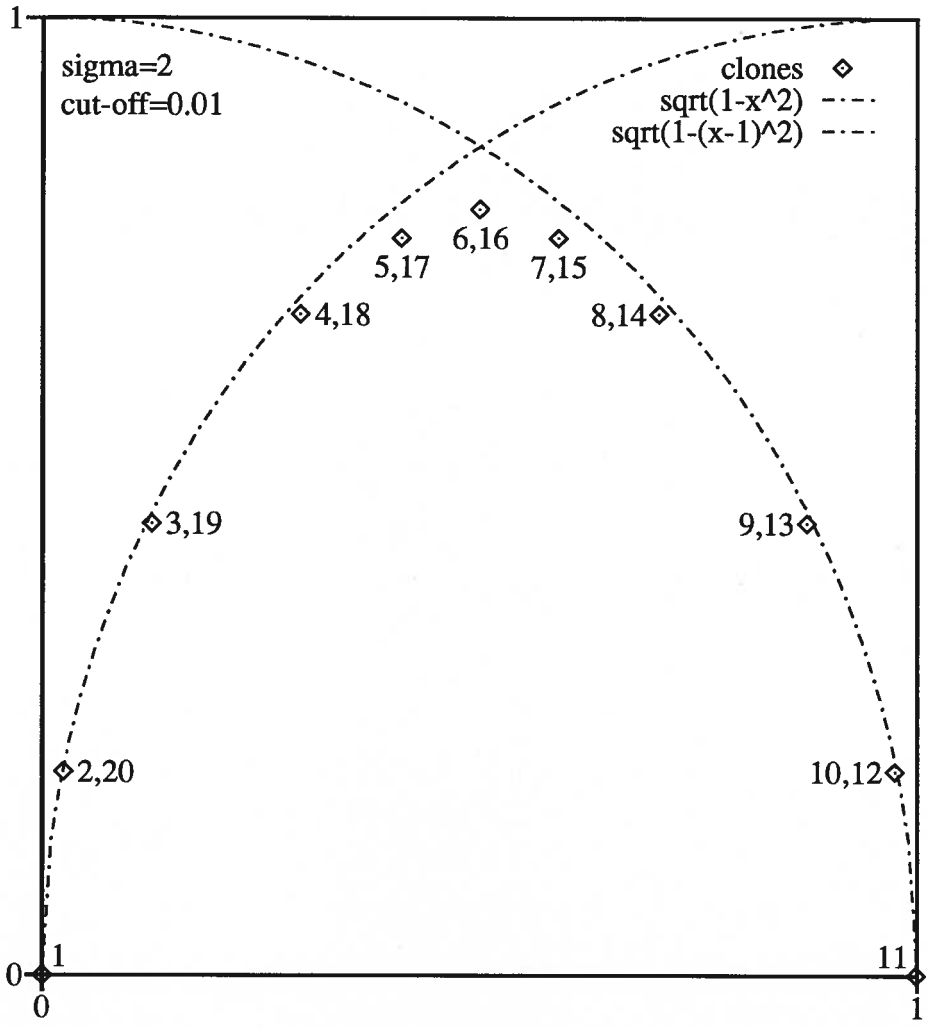

Figure 6.47: The distance coefficient representation for the periodic Δ shape-space, with uniform initial conditions. The positions of the clones are fixed in time and are symmetrically positioned with respect to complementary reference clones 1 and 11. $x_i = 0.1$, $S = 10$, $k_2 = 1$, $k_3 = 0.01$, $k_4 = 1$, $C_2 = 0.6$, $C_3 = 0.7$, $\sigma = 2$, *cut-off*$= 10^{-3}$.

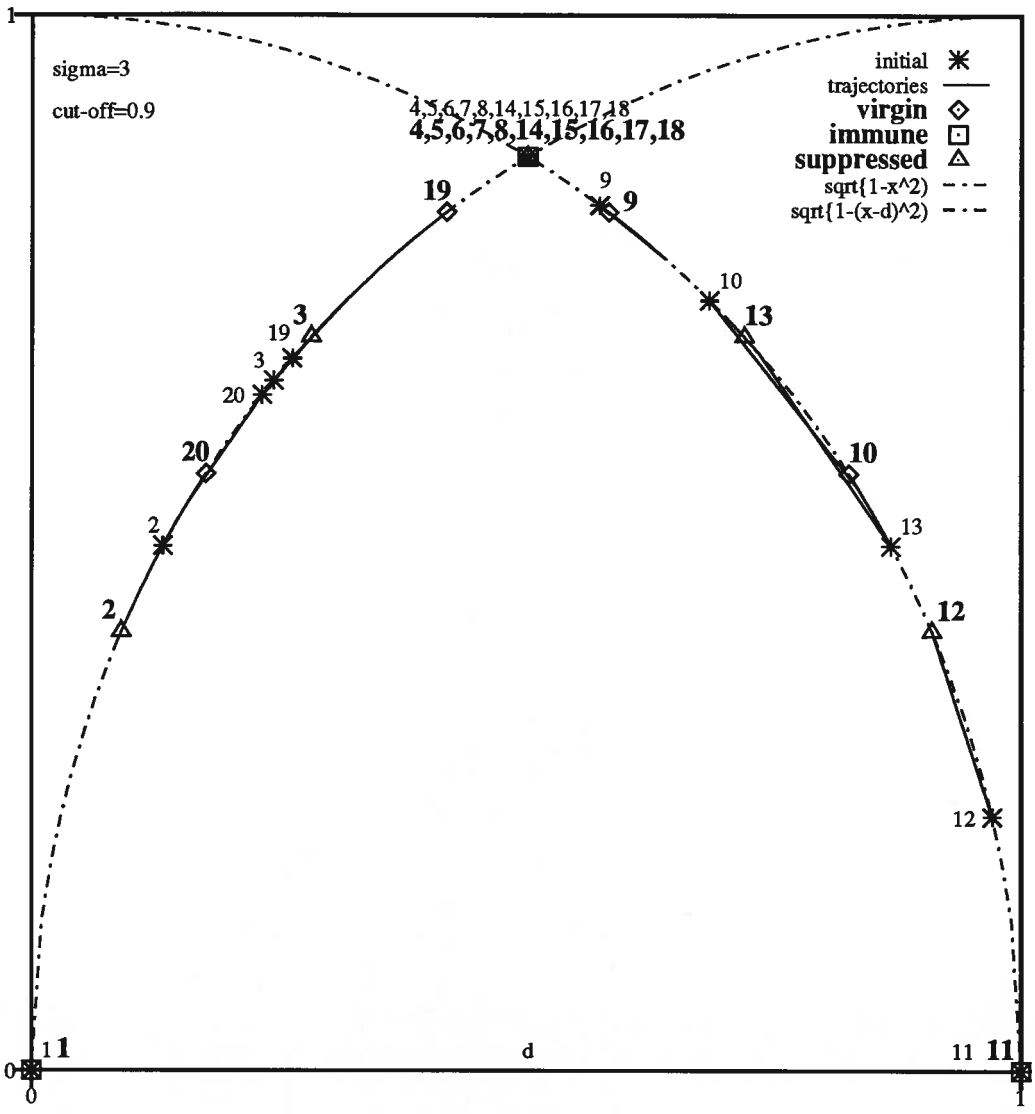

Figure 6.48: The distance coefficient representation for the periodic Δ shape-space of a 20-dimensional network with random initial conditions. d is the distance between reference clones 1 and 11. $S = 10$, $k_2 = 1$, $k_3 = 0.01$, $k_4 = 1$. $\sigma = 3$, *cut-off*= 0.9, $C_2 = 0.45$, $C_3 = 0.65$.

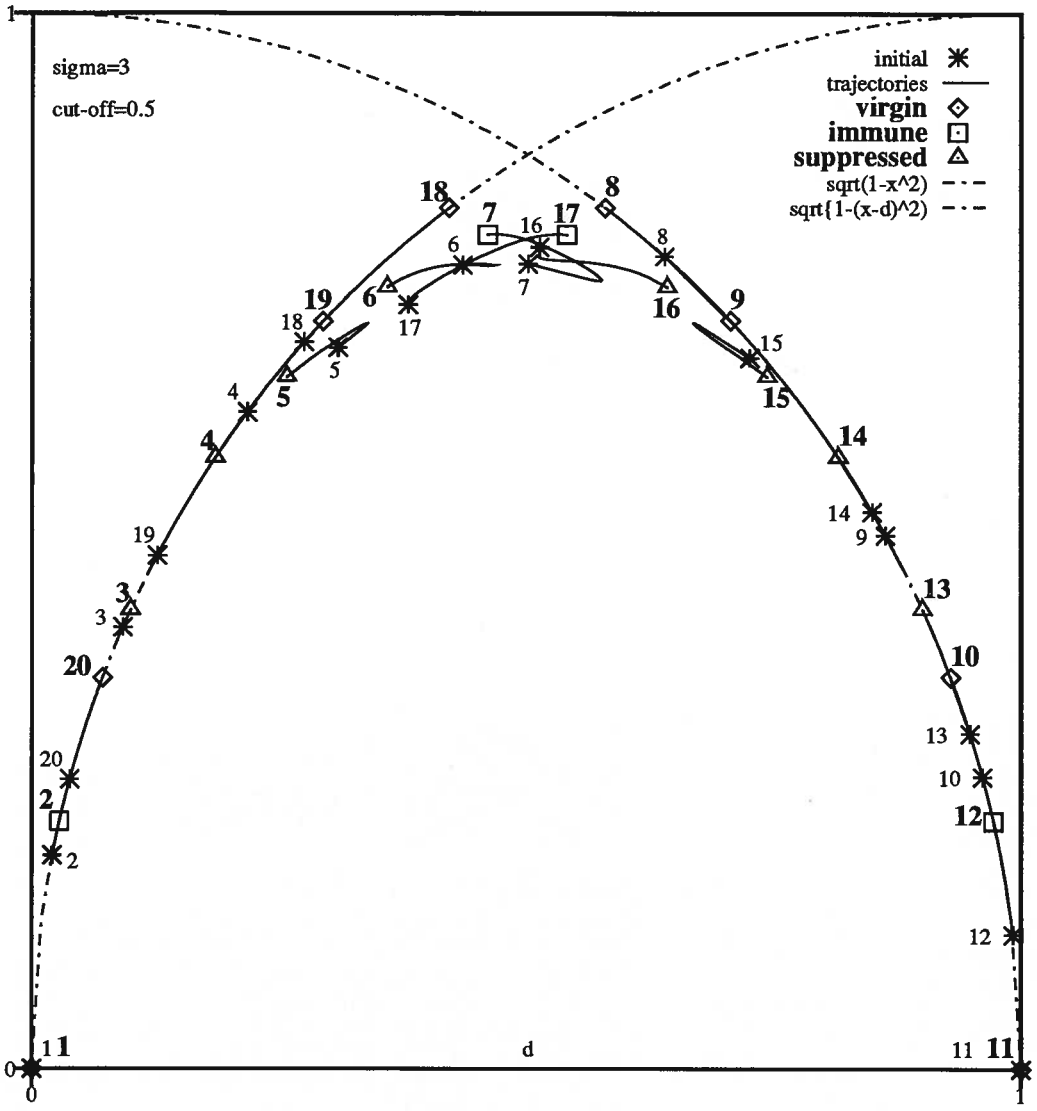

Figure 6.49: Same as in figure 6.48 but $\sigma = 3$, *cut-off*= 0.5, $C_2 = 0.65$, $C_3 = 0.7$.

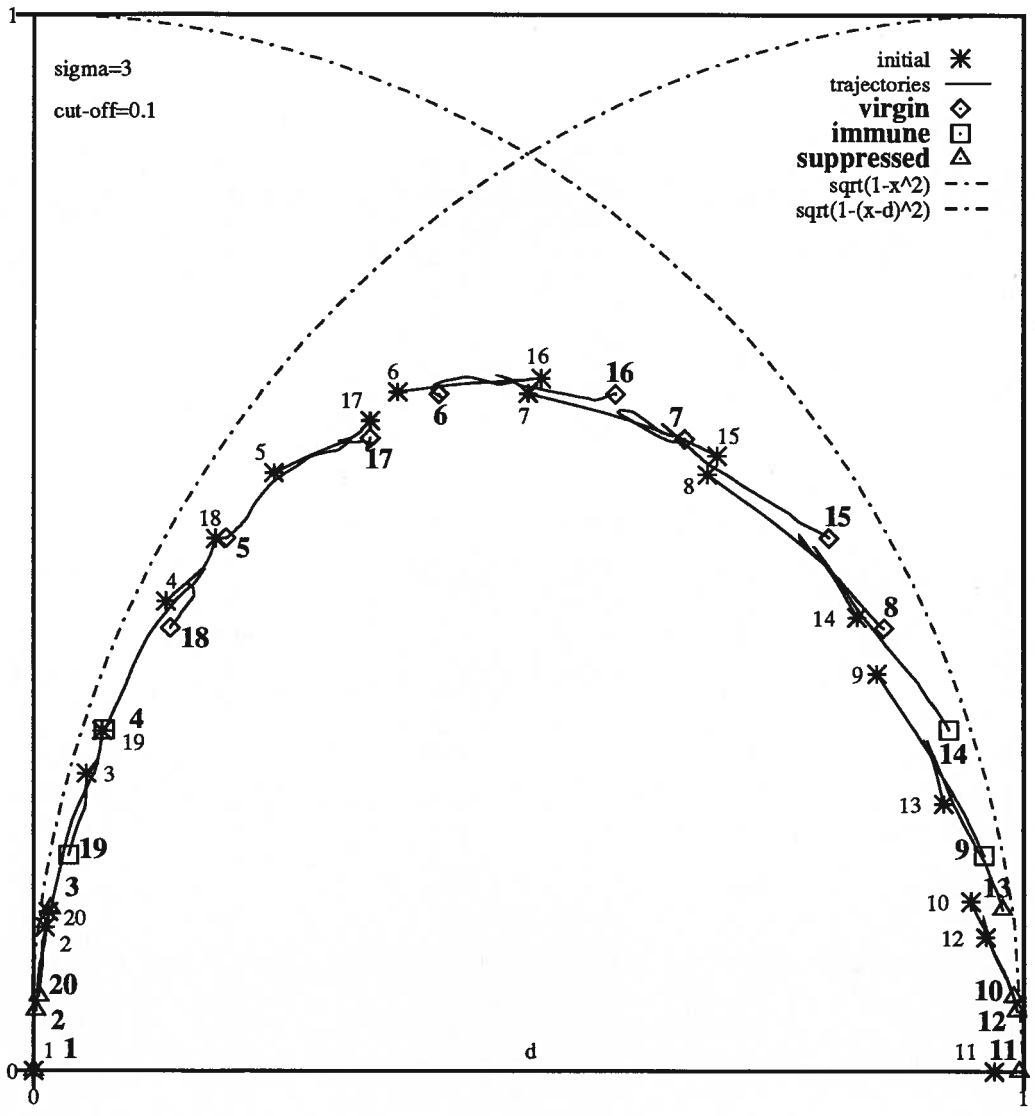

Figure 6.50: Same as in figure 6.48 but $\sigma = 3$, *cut-off*= 0.1, $C_2 = 0.65$, $C_3 = 0.7$.

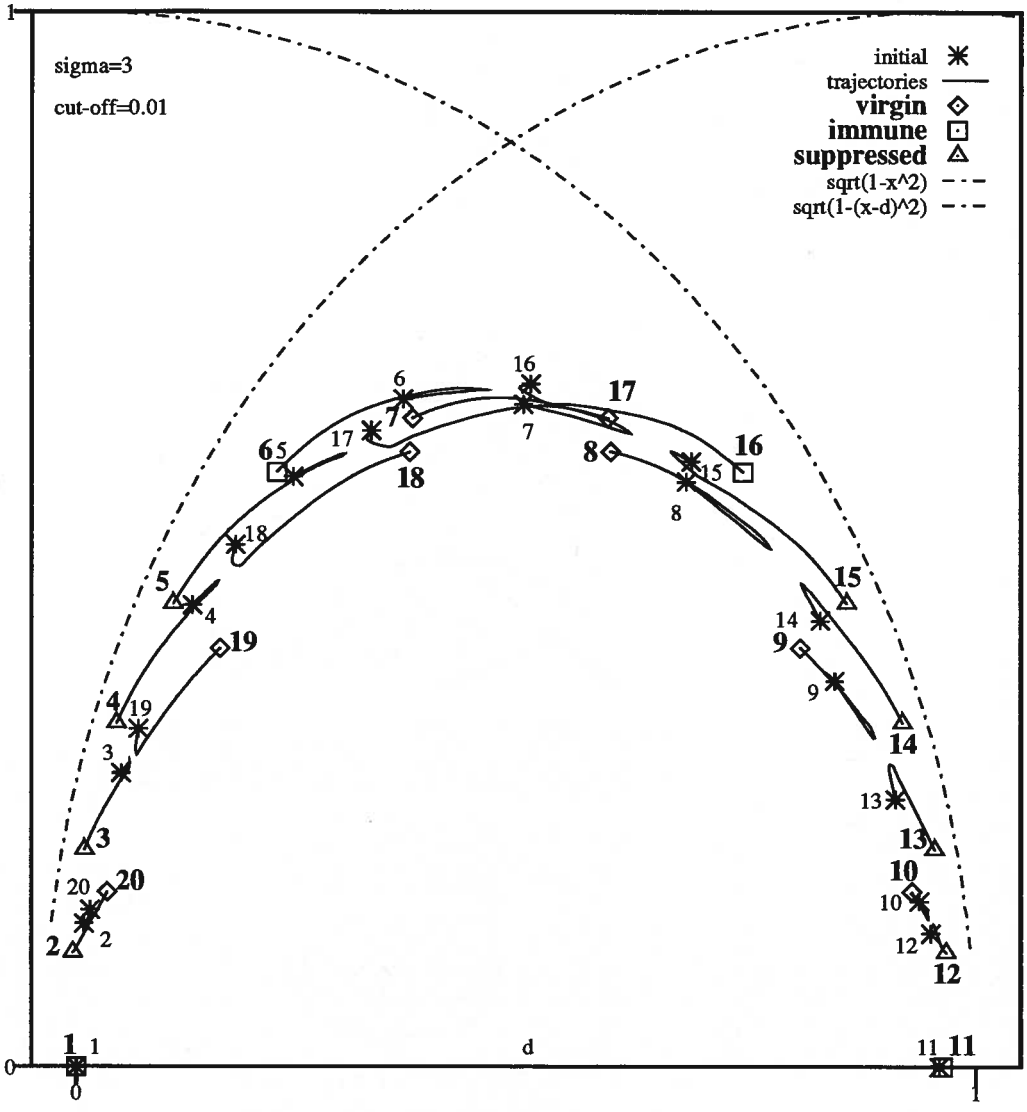

Figure 6.51: Same as in figure 6.48 but $\sigma = 3$, *cut-off*= 0.01, $C_2 = 0.65$, $C_3 = 0.68$.

the three different symbols that identifies their final state: virgin, immune or suppressed. The dashed curves are part of two circles of centers equal to the final positions of the reference clones 1 and 11, and of radius 1 (the formula that generates the curve for clone 11 contains the term d, which is equal to the distance between the final positions of clones 1 and 11).

It is observed that the similarity computed by the distance coefficient is here also quite in accordance with the one which is implicit in the periodic Δ shape-space. For example, clones 2 and 20 are neighbours of clone 1 in shape-space. They also are in the distance coefficient representation. Also, there seems to be a tendency for suppressed clones to move less in the distance coefficient plane than the virgin or immune clones.

It can be seen that for a wide Gaussian (which is the case in figure 6.48 where *cut-off=0.01*, the dynamics are happening along a wide arc. But for a narrower Gaussian (for example, in figure 6.49, the *cut-off* is 0.1), the dynamics happen along a narrower band.

The system was also studied for a smaller value of the variance, that is $\sigma = 2$. It was found that the distances between clones are then bigger (not shown).

6.7.2 Random boolean affinities

One could easily observe that in the previous cases where one used affinities computed in relation to a one-dimensional shape-space, the positions of the clones in the distance coefficients were observing a somewhat smooth pattern along a more or less wide semi-arc. One can imagine that this regular pattern arises from the particular choice of the affinities. To test this hypothesis, one can make a simulation with random Boolean affinities. The matrix of affinities is shown in figure 6.52. The simulation was done for a 22-dimensional network and random initial conditions, with random Boolean affinities. It is shown in figure 6.53 and does indeed verify that the previous regular pattern arose

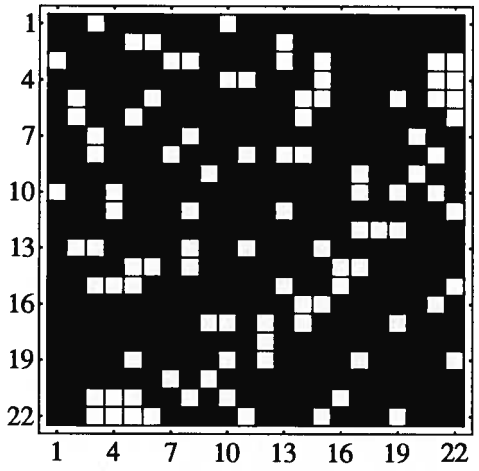

Figure 6.52: The two-dimensional representation of a matrix of affinities for a 22-dimensional network, with random Boolean affinities. Clones are numbered from 1 to 22. In the two-dimensional representation, the squares shaded white represent an afffinity value of one and the ones shaded black represent an affinity value of zero.

from the particular choice of the affinities computed in the shape-space context. It can also be noticed that in the last simulation of the distance coefficient using affinities from the shape-space, that is with $\sigma = 0.9$, there were many clones that would be considered as maximally dissimilar from reference clones 1 and 11. This arose from the fact that the clones far away from the reference clones in shape-space were not very much connected to other clones. In the random Boolean case, depending on the connectance (percentage of ones in the matrix), one can have more or less clones maximally distant from the reference clones. For a smaller connectance, most clones are not very much connected with others, and thus there is a great probability that they would be considered as maximally dissimilar from other clones, including the reference clones. As one increases the connectance, there are more connections established. This improves the chances of having clones being more similar from the point of view of the network.

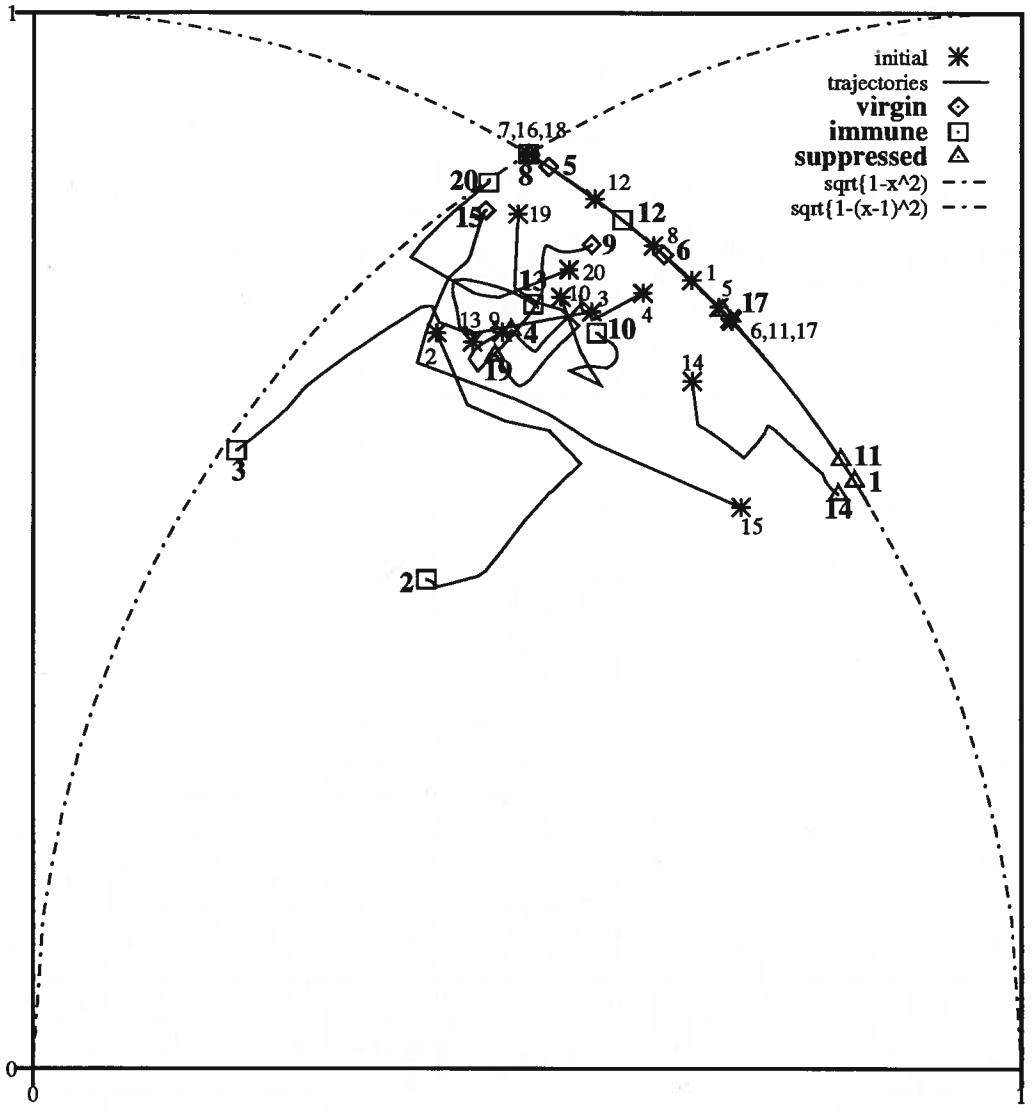

Figure 6.53: The distance coefficient representation for a 22-dimensional network of clones randomly connected with Boolean affinities. One can see that the regular pattern that was exhibited for the affinities derived from a one-dimensional shape-space is not present anymore. $0 < x_i < 1$, $S = 10$, $k_2 = 1$, $k_3 = 10$, $k_4 = 1$, $C_2 = 10$, $C_3 = 3$.

So the similarity between this simulation and the one done in shape-space is that as one increases the connectance of the matrix or the variance of the Gaussian, one gets closer to the mean-field (all clones are connected to each other in the same way). And then they should all appear to be equally similar (this would be the case if the concentrations of the clones were all identical). In the other extreme case, one could have a situation where there is an almost null connectance or a very narrow variance of the Gaussian. And then almost all clones would be considered as greatly dissimilar to each other.

This ends the work presented in this thesis.

Chapter 7

Conclusion

This thesis has covered a wide range of subjects. It started by recalling the first observations of immunity which gave rise to the theories of acquired immunity, then presented the theories of antibody formation which culminated in the theory of clonal selection. This was important so that the reader had some background for the understanding of the immune network hypothesis which pushed so many theorists to show interest in modelling the immune network. Hoffmann's N-dimensional model was then presented. It was said that this model had never been studied with affinities other than Boolean. So some non-Boolean possible matrices were presented. A particular one based on the shape-space formulated by Segel and Perelson was chosen. This led to the creation of a new shape-space without shape zero. Some simulations and analysis of Hoffmann's N-dimensional network model with non-Boolean affinities derived from those two versions of shape-space were shown. And finally the main part could be presented, that is the review, analysis and generalization of the distance coefficient in terms of non-Boolean affinities, as well as numerical simulations of the distance coefficients of clones of an immune network modelled by Hoffmann's N-dimensional network model.

A summary of the new material that was presented in this thesis can be made. It consists of implementing non-Boolean affinities in Hoffmann's model, more specifically affinities from a shape-space. Also, a new shape-space setting is presented and simulations with various mappings onto a plane of the N-dimensional dynamics are done with both the original and the new shape-space. A mapping using distance coefficients is presented.

The distance coefficient is analyzed and generalized for its use with non-Boolean affinities. Simulations are done with both Boolean and non-Boolean affinities.

However, as it was mentioned several times already, because the work presented was broad it was not very deep to keep the length of the thesis within reasonable bounds. It thus calls for further work, that is analysis and simulations. It is hoped, though, that it aroused the curiosity of the reader and showed what interesting work can be performed in theoretical immunology.

Appendix A

The N-dimensional model brought back to a 1-dimensional model

It is here shown that the N-dimensional system behaves like a 1-dimensional one when periodic boundary conditions on a shape-space and uniform initial clone populations are considered.

Consider the field of a clone k, given by expression (3.6). Now, due to the specific initial conditions, the concentration x_k is identical for each clone. It can be characterized by a constant x and factorized in front of the summation. Furthermore, because of the periodicity the shape-space, the sum of the affinities relating a clone to the others in the system will be the same. As a consequence, the field experienced by each clone is the same and can be simply written as:

$$Y = x\,K \tag{A.72}$$

where it is understood that K is the summation over the affinities of one clone with the others.

The same reasoning allows to show that the parameter expressed in (3.7) that gives a measure of the T-cell factors are identical for all clones and given by:

$$W = x\,Y\,K. \tag{A.73}$$

Note that in this case it would not make sense to normalize W. See what happens if one wants to do so. Remember that in the original formulation of the model presented by Hoffmann et al. [61], one first calculates the T-cell factor concentration for each clone

and then multiplies it by a factor equal to the average of the two threshold factors C_2 and C_3 divided by the average of all the T-cell factors concentrations. Now examine what one gets when applying the normalization to this case.

The average of the T-cell factors concentrations is here simply given by W. The normalized value of the T-cell factors concentration is then given by:

$$\begin{aligned} W' &= W\,\frac{(C_2+C_3)/2}{W} \\ &= \frac{C_2+C_3}{2}. \end{aligned} \tag{A.74}$$

which is a constant! It is thus independant of the change in concentration of the clones, which does not make sense. It is concluded that the the normalization of the T-cell factors concentrations is not required when dealing with uniform initial conditions.

The next step is to consider the effectivities, given by (3.5). As they only depend on the T-cell factors concentrations, one can see that they are also identical for all clones:

$$e_q = \frac{1}{1+\left(\dfrac{W}{C_q}\right)^{n_q}} \qquad q=1,2,3. \tag{A.75}$$

This last finding leads to the fact that the effective fields expressed by (3.9) are also identical for all clones and simply given by:

$$U_q = x\,e_q\,K \qquad q=1,2,3. \tag{A.76}$$

Finally, the basic differential equation (3.4), in respect of what has just been shown, can be written as:

$$\frac{dx}{dt} = S + k_1\,x\,e_1\,U_1 - k_2\,x\,e_2\,U_2 - k_3\,x\,e_3\,(U_3)^2 - k_4\,x. \tag{A.77}$$

It can be seen that the same equation is controlling the behaviour of the uniformly distributed initial populations of the cells, and therefore these populations will all vary

identically. For that reason, the model can be seen as one-dimensional in this special situation and will later be refered as so. Well, this is not a very interesting case in itself, but one of the advantages of its extreme simplicity is to enable the calculation of parameters that yield biologically reasonable steady-states. These parameters can then be used for further investigations of the model.

To find them, one can borrow the same idea that was used previously in the first of paper on Hoffmann's symmetrical two-dimensional models [57]. It is an approximate method of analysis that allows to find parameters that generate predetermined steady-states. One considers each of the steady-states to be essentially a balance between two dominant terms in the differential equation. Here, the case where $C_2 < C_3$ is studied and $k_1 = 0$.

In this approach, some biological assumptions are made. First, the virgin state x_v is characterized by a balance between source and IgM killing terms:

$$S \approx k_2 \, x_v \, U_2 \, e_2. \tag{A.78}$$

This can be solved to yield, in view of (A.76):

$$x_v \approx \left(\frac{S}{k_2 \, K} \right)^{1/2}. \tag{A.79}$$

This solution can occur only if $W_v < C_2$, i.e. if:

$$\frac{S \, K}{k_2} < C_2. \tag{A.80}$$

The immune state x_i, represented by a balance between source and IgG killing terms is given by:

$$S \approx k_3 \, x_i \, (U_3)^2 \, e_3. \tag{A.81}$$

This yields:

$$x_i \approx \left(\frac{S}{k_3 \, K^2} \right)^{1/3}. \tag{A.82}$$

This result requires that $C_2 < W_i < C_3$. So one must have:

$$C_2 < \left(\frac{S\,K}{k_3}\right)^{2/3} < C_3. \tag{A.83}$$

Finally, the suppressed state x_s is a balance between source and natural death terms:

$$S \approx k_4\, x_s. \tag{A.84}$$

This yields:

$$x_s \approx \frac{S}{k_4}. \tag{A.85}$$

This gives the condition that $W_s > C_3$, i.e. :

$$\left(\frac{S\,K}{k_4}\right)^{2} > C_3. \tag{A.86}$$

Combining the above inequalities gives the conditions on the parameters which yield those steady-states:

$$\frac{S\,K}{k_2} < C_2 < \left(\frac{S\,K}{k_3}\right)^{2/3} < C_3 < \left(\frac{S\,K}{k_4}\right)^{2}. \tag{A.87}$$

A similar procedure was used by Hoffmann et al. [61] when they considered their one-dimensional model using Boolean affinities. It appears that the only difference between inequality (A.87) and the one they find is the presence of the sum of the affinities in the first case!

Appendix B

Complementary clones can have the same concentration

B.1 The periodic Δ shape-space

It is shown here that for the periodic Δ shape-space, a solution that satisfies the N-dimensional model exists for which complementary clones have the same concentration. An important feature that makes this solution possible appears to be the periodicity associated with the affinities which allows to equate the expressions for the steady-states of the two complementary clones. Indeed, the following relationships exist, at least, between the affinities:

$$K_{ij} = K_{ji} = K_{nm} = K_{mn} \tag{B.88}$$

$$\begin{aligned} \text{where} \quad i,j &= 1,N \\ m &= \begin{cases} i + N/2 & , i \leq N/2 \\ i - N/2 & , i > N/2 \end{cases} \\ n &= \begin{cases} j + N/2 & , j \leq N/2 \\ j - N/2 & , j > N/2. \end{cases} \end{aligned} \tag{B.89}$$

The fields of two complementary clones k and $k + N/2$, where $k = 1, N/2$ are given accordingly to (3.6) by:

$$\begin{aligned} Y_k &= K_{k,1}\, x_1 + \ldots + K_{k,N/2}\, x_{N/2} \\ &\quad + K_{k,1+N/2}\, x_{1+N/2} + \ldots + K_{k,N}\, x_N \\ Y_{k+N/2} &= K_{k+N/2,1}\, x_1 + \ldots + K_{k+N/2,N/2}\, x_{N/2} \end{aligned} \tag{B.90}$$

$$+ K_{k+N/2,1+N/2}\, x_{1+N/2} + \ldots + K_{k+N/2,N}\, x_N \tag{B.91}$$

where $k = 1, N/2$. But the relationships (B.89) between the affinities allow to write that:

$$K_{k,1} = K_{k+N/2,1+N/2}; \ldots; K_{k,N/2} = K_{k+N/2,N} \tag{B.92}$$

$$K_{k,1+N/2} = K_{k+N/2,1}; \ldots; K_{k,N} = K_{k+N/2,N/2}. \tag{B.93}$$

We can then rewrite (B.91), using these relationships, as:

$$\begin{aligned} Y_{k+N/2} &= K_{k,1}\, x_{1+N/2} + \ldots + K_{k,N/2}\, x_N \\ &\quad + K_{k,1+N/2}\, x_1 + \ldots + K_{k,N}\, x_{N/2} \end{aligned} \tag{B.94}$$

with the order of the terms in the summation of (B.91) reversed. Now assume that complementary shapes that have the same concentration is a solution. One can then replace each cell concentration by the concentration of the complementary clone, that is, one replaces x_i by $x_{i+N/2}$ for $i \leq N/2$ and x_i by $x_{i-N/2}$ for $i > n/2$. It can easily be seen that one then has that $Y_k = Y_{k+N/2}$.

The T-cell factor concentrations of the same two complementary clones k and $k+N/2$ are given, accordingly to (3.7), by:

$$\begin{aligned} W_k &= K_{k,1}\, x_1\, Y_1 + \ldots + K_{k,N/2}\, x_{N/2}\, Y_{N/2} \\ &\quad + K_{k,1+N/2}\, x_{1+N/2}\, Y_{1+N/2} + \ldots + K_{k,N}\, x_N\, Y_N \end{aligned} \tag{B.95}$$

$$\begin{aligned} W_{k+N/2} &= K_{k+N/2,1}\, x_1\, Y_1 + \ldots + K_{k+N/2,N/2}\, x_{N/2}\, Y_{N/2} \\ &\quad + K_{k+N/2,1+N/2}\, x_{1+N/2}\, Y_{1+N/2} + \ldots + K_{k+N/2,N}\, x_N\, Y_N. \end{aligned} \tag{B.96}$$

There is no need to consider the normalized T-cell factor concentrations since the presence of the normalization constant does not invalid the present verification. The expression for $W_{k+N/2}$ can be rewritten in a similar way that was done for $Y_{k+N/2}$, as:

$$\begin{aligned} W_{k+N/2} &= K_{k,1}\, x_{1+N/2}\, Y_{1+N/2} + \ldots + K_{k,N/2}\, x_N\, Y_N \\ &\quad + K_{k,1+N/2}\, x_1\, Y_1 + \ldots + K_{k,N}\, x_{N/2}\, Y_{N/2}. \end{aligned} \tag{B.97}$$

Again, assuming that complementary clones have the same concentration shows that, their fields also having the same values as shown above, their T-cell concentrations are the same, that is $W_k = W_{k+N/2}$. This also means that $e_k = e_{k+N/2}$, since e_k is directly proportional to W_k, according to (3.5).

The effective fields of the two complementary clones k and $k+N/2$ are given, accordingly to (3.9), by:

$$\begin{aligned} U_k &= K_{k,1}\, e_1\, x_1 + \ldots + K_{k,N/2}\, e_{N/2}\, x_{N/2} \\ &\quad + K_{k,1+N/2}\, e_{1+N/2}\, x_{1+N/2} + \ldots + K_{k,N}\, e_N\, x_N \qquad \text{(B.98)} \\ U_{k+N/2} &= K_{k+N/2,1}\, e_1\, x_1 + \ldots + K_{k+N/2,N/2}\, e_{N/2}\, x_{N/2} \\ &\quad + K_{k+N/2,1+N/2}\, x_{1+N/2} + \ldots + K_{k+N/2,N}\, x_N. \qquad \text{(B.99)} \end{aligned}$$

Again, the same work can be performed on the expression for $U_{k+N/2}$ to yield the new expression for (B.99):

$$\begin{aligned} U_{k+N/2} &= K_{k,1}\, e_{1+N/2}\, x_{1+N/2} + \ldots + K_{k,N/2}\, e_{1+N/2}\, x_N \\ &\quad + K_{k,1+N/2}\, e_1\, x_1 + \ldots + K_{k,N}\ e_{N/2}\, x_{N/2}. \qquad \text{(B.100)} \end{aligned}$$

It was already said above that assuming complementary clones to have the same value makes the effectivities have the same value, and as a direct consequence the effective fields also are seen to take the same value.

So, looking back at the expressions (3.18) and (3.20) that describe respectively the virgin and immune states, one can see that complementary clones that have the same concentration and also have the same effective fields is a solution that satisfies those expressions. It is obvious that in the suppressed state complementary clones that share the same concentration is a solution, since its value is unique.

B.2 The non-periodic ± shape-space

It is shown here that for the non-periodic shape-space, there exists a solution that satisfies the N-dimensional network for which complementary clones have the same concentration. One here numbers the clones from $-N$ to N, where N is any number. The fields of two complementary clones k and $-k$ are given accordingly to (3.6), by:

$$Y_k = K_{k,-N}\, x_{-N} + \ldots + K_{k,0}\, x_0 + \ldots + K_{k,N}\, x_N \tag{B.101}$$

$$Y_{-k} = K_{-k,-N}\, x_{-N} + \ldots + K_{-k,0}\, x_0 + \ldots + K_{-k,N}\, x_N. \tag{B.102}$$

But the following relationships exist between the affinities:

$$K_{k,-N} = K_{-k,N}, \ldots, K_{k,0} = K_{-k,0}, \ldots, K_{k,N} = K_{-k,-N}. \tag{B.103}$$

By doing the similar type of work that was done in Appendix B.1, on can rewrite (B.102) as:

$$Y_{-k} = K_{k,-N}\, x_N + \ldots + K_{k,0}\, x_0 + \ldots + K_{k,N}\, x_{-N}. \tag{B.104}$$

Assuming that complementary clones having the same concentration (i.e. $x_k = x_{-k}$ for $k = -N, N$) is a solution shows that in that case their fields take the same value, that is: $Y_k = Y_{-k}$ for $k = -N, N$.

The T-cell factors concentrations of two complementary clones k and $-k$ are given by:

$$W_i = K_{i,-N}\, x_{-N}\, Y_{-N} + \ldots + K_{i,0}\, x_0\, Y_{-N} + \ldots + K_{i,N}\, x_N\, Y_N \tag{B.105}$$

$$W_{-i} = K_{-i,-N}\, x_{-N}\, Y_{-N} + \ldots + K_{-i,0}\, x_0\, Y_0 + \ldots + K_{-i,N}\, x_N\, Y_N. \tag{B.106}$$

This last expression can be written using relationships (B.103) as:

$$W_{-i} = K_{i,-N}\, x_N\, Y_N + \ldots + K_{i,0}\, x_0\, Y_{-N} + \ldots + K_{i,N}\, x_{-N}\, Y_{-N}. \tag{B.107}$$

If complementary clones have the same concentration, then their fields have been shown to have the same value, and therefore the T-cell factor concentrations also have the same value, that is: $W_k = W_{-k}$ for $k = -N, N$. Again, since the effectivities depend only on the T-cell factors concentrations, they also take the same value in that case.

The effective fields of two complementary clones k and $-k$ are given by:

$$U_i = K_{i,-N}\, e_{-N}\, x_{-N} + \ldots + K_{i,0}\, e_0\, x_0 + \ldots + K_{i,N}\, e_N\, x_N \tag{B.108}$$

$$U_{-i} = K_{-i,-N}\, e_{-N}\, x_{-N} + \ldots + K_{-i,0}\, e_0\, x_0 + \ldots + K_{-i,N}\, e_N\, x_N. \tag{B.109}$$

This last expression can be rewritten using relationships (B.103) to yield:

$$U_i = K_{i,-N}\, e_N\, x_N + \ldots + K_{i,0}\, e_0\, x_0 + \ldots + K_{i,N}\, e_{-N}\, x_{-N}. \tag{B.110}$$

The effective fields of complementary clones are easily shown to take the same value if complementary clones that have the same concentration is assumed to be a solution, since then their effectivities have been shown to take the same value. One then has that $U_k = U_{-k}$ for $k = -N, N$.

So looking back at expressions (3.18) and (3.20) that describe respectively the virgin and the immune states, one can see that complementary clones that have the same concentration and also have the same effective fields is a solution that satisfies those expressions. The same thing can be said for the suppressed state for the same reasons that were stated in Appendix B.1.

Appendix C

The distance coefficients of Hoffmann and Tufaro form a metric

The following is essentially a slightly reworded review of the demonstration done by Hoffmann and Tufaro [62] that serological distance coefficients form a metric. Consider any three sera X, Y and Z, and a reagent C. In order for the distance coefficients between these sera to form a metric, they must obey the following criteria [83]:

(1) $D[\mathrm{X,Y|C}] \geq 0$

(2) $D[\mathrm{X,X|C}] = 0$

(3) $D[\mathrm{X,Y|C}] = D[\mathrm{Y,X|C}]$

(4) If $\mathrm{X} \neq \mathrm{Y}$, then $D[\mathrm{X,Y|C}] > 0$

(5) $D[\mathrm{X,Y|C}] \leq D[\mathrm{Y,Z|C}] + D[\mathrm{Z,X|C}]$.

Using Venn diagrams, one can denote the fractions of C that react with different groups of sera the following way:

$x \equiv$ the fraction of C that reacts with X only

$y \equiv$ the fraction of C that reacts with Y only

$z \equiv$ the fraction of C that reacts with Z only

$u \equiv$ the fraction of C that reacts with Y and Z only

$v \equiv$ the fraction of C that reacts with X and Z only

$w \equiv$ the fraction of C that reacts with X and Y only

$t \equiv$ the fraction of C that reacts with X, Y and Z

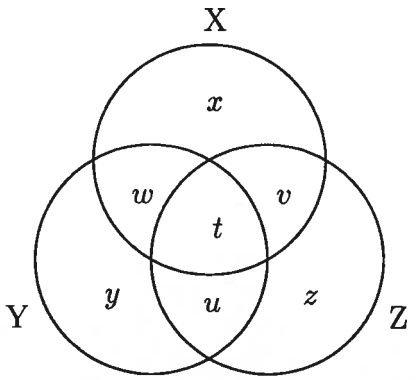

Figure C.54: The Venn diagrams for the fractions of C that interact with three sera X, Y and Z. (Adapted from [62].)

$$\text{where } x, y, z, u, v, w, t \geq 0. \tag{C.111}$$

With respect to these, the distance coefficient between sera X and Y is given[1] by:

$$D[\text{X,Y}|\text{C}] = \frac{x + v + y + u}{x + v + y + u + w + t}. \tag{C.112}$$

The distance coefficients between sera Y and Z, and between sera X and Y can be written in a similar way.

The proof that the distance coefficients obey the metric criteria follows immediately.

(1) The fact that the denominator of (C.112) contains all the terms in the numerator plus some others, combined with (C.111) automatically leads to the conclusion that criterion (1) is obeyed.

(2) For two sera X and Y where $X = Y$, one has that $x = v = y = u = 0$. So one gets that $D[\text{X,X}|\text{C}] = 0$, which satisfies criterion (2).

(3) The expression giving $D[\text{Y,X}|\text{C}]$ is identical to the one giving $D[\text{X,Y}|\text{C}]$, which satisfies criterion (3).

[1] using the first definition 6.31

(4) First note that D[X,Y|C] can be zero even if $X \neq Y$, since they are identical in the eyes of C if they interact identically with it. This can seem to violate criterion (4). But one has to realize that one does only consider distances *from the perspective of C.* With this regard, sera X and Y are identical from the perspective of C if they interact identically with C and then D[X,Y|C] = 0, which satisfies criterion (2). But if they interact differently with C, then they are different from the perspective of C and then D[X,Y|C] > 0, which satisfies criterion (4).

(5) It needs to be shown that:

$$D[\mathrm{X},\mathrm{Y}|\mathrm{C}] + D[\mathrm{Y},\mathrm{Z}|\mathrm{C}] \geq D[\mathrm{X},\mathrm{Z}|\mathrm{C}] \tag{C.113}$$

which is equivalent to asking that:

$$\frac{x+v+y+u}{x+v+y+u+w+t} + \frac{x+w+z+u}{x+w+z+u+v+t} \geq \frac{y+w+z+v}{y+w+z+v+u+t} \tag{C.114}$$

This last inequality can be rewritten as:

$$\begin{aligned} & (x+v+y+u)(x+w+z+u+v+t)(y+w+z+v+u+t) \\ + \quad & (x+w+z+u)(x+v+y+u+w+t)(y+w+z+v+u+t) \\ \geq \quad & (y+w+z+v)(x+v+y+u+w+t)(x+w+z+u+v+t) \end{aligned} \tag{C.115}$$

Since the left-hand side of this inequality contains all the terms contained in the right-hand side plus others, and because of (C.111), the inequality is respected. This satisfies criterion (5).

Bibliography

[1] G. Adam. Theoretical models of lymphocyte network interactions in ontogenic generation of antibody diversity. In G. I. Bell, A. S. Perelson, and G. H. Pimbley, editors, *Theoretical Immunology*, pages 603–627. Marcel Dekker, New York, 1978.

[2] G. Adam and E. Weiler. Lymphocyte population dynamics during ontogenetic generation of diversity. In A. J. Cunningham, editor, *The generation of antibody diversity.* Academic Press, 1976.

[3] J. Alexander. *Protoplasma*, 14(296), 1931. (quoted in [125], page 69).

[4] O. Bail. *Z. Immunitaetsforsch.*, 21(202), 1914. (quoted in [125], page 67).

[5] O. Bail and H. Rotky. *Z. Immunitaetsforsch.*, 17(378), 1913. (quoted in [125], page 67).

[6] O. Bail and K. Tsuda. *Z. Immunitaetsforsch.*, 1(546, 772), 1909. (quoted in [125], page 67).

[7] G. I. Bell. Mathematical model of clonal selection and antibody production,I. *J. Theor. Biol.*, 29:191–232, 1970.

[8] G. I. Bell. Mathematical model of clonal selection and antibody production,II. *J. Theor. Biol.*, 33:339–378, 1971.

[9] G. I. Bell. Mathematical model of clonal selection and antibody production,III. *J. Theor. Biol.*, 33:379–398, 1971.

[10] E. Berger and H. Erlenmeyer. *Z. Hyg. Infektionskr.*, 113(79), 1931. (quoted in [125], page 64).

[11] A. Besredka. *Ann. Inst. Pasteur, Paris*, 15(785), 1901. (quoted in [125], page 263).

[12] J. Bordet. *Ann. Inst. Pasteur, Paris*, 12(688), 1899. (quoted in [125], page 51).

[13] J. Bordet. *Ann. Inst. Pasteur, Paris*, 13(273), 1899. (quoted in [125], page 263).

[14] J. Bordet. *Ann. Inst. Pasteur, Paris*, 14(257), 1900.

[15] J. D. Bowman and F. Aladjem. A method for the determination of heterogeneity of antibodies. *J. Theor. Biol.*, 4:242–253, 1963.

[16] F. Breinl and F. Haurowitz. *Z. Physiol. Chem.*, 192(45), 1930. (quoted in [125], page 69).

[17] C. Bruni, A. Gandolfi, and A. Germani. Theoretical investigation and numerical procedure for the estimation of antibocy distribution from binding data. In G. I. Marchuk and L. N. Belykh, editors, *Mathematical Modeling in Immunology and Medicine*, pages 97–108, New-York, 1983. IFIP, North-Holland.

[18] C. Bruni, G. Koch, and R. Strom. Derivation of antibody distribution from experimental binding data. *J. Theor. Biol.*, 61:143–170, 1979.

[19] H. Buchner. *Muench. Med. Wochenschr.*, 40(449, 480, 482), August 1893. (quoted in [125], page 61).

[20] F. M. Burnet. *The production of antibodies.* Macmillan, New York, 1941. (quoted in [125], page 72).

[21] F. M. Burnet. *Aust. J. Sci.*, 20(67), 1957. (quoted in [125], page 79).

[22] F. M. Burnet. *The Clonal Selection Theory of Acquired Immunity.* Cambridge University Press, London, 1959. (quoted in [125], page 79).

[23] F. M. Burnet and F. Fenner. *The production of antibodies.* Macmillan, New York, 2nd edition, 1949. (quoted in [125], page 73).

[24] L. Camus and E. Gley. *Ann. Inst. Pasteur, Paris*, 13(779), 1899. (quoted in [125], page 262).

[25] F. H. C. Crick. *Symp. Soc. Exp. Biol.*, 12(138), 1958. (quoted in [125], page 78).

[26] S. Dasgupta. *Physica A*, 403:189, 1992.

[27] R. J. De Boer. Symmetric idiotypic networks: connectance and switching, stability and suppression. In A. S. Perelson, editor, *Theoretical immunology, Part Two*, pages 265–289. Addison-Wesley, Redwood City, CA., 1988.

[28] R. J. De Boer. Extensive percolation in reasonable idiotypic networks. In H. Atlan and I. R. Cohen, editors, *Theories of immune networks*, volume 46 of *Springer Series in Synergetics*, Heidelberg, 1989. Springer-Verlag Berlin.

[29] R. J. De Boer and P. Hogeweg. Memory but no suppresssion in low-dimensional symmetric idiotypic networks. *Bull. Math. Biol*, 51, 1989.

[30] R. J. De Boer and P. Hogeweg. Unreasonable implications of reasonable idiotypic network assumptions. *Bull. Math. Biol.*, 51:381–408, 1989.

[31] R. J. De Boer, I. G. Kevrekedis, and A. S. Perelson. A simple idiotypic network model with complex dynamics. *Chem. Eng. Sci.*, 45:2375–2382, 1990.

[32] R. J. De Boer and A. S. Perelson. Size and connectivity as emergent properties of a developing immune system. *J. Theor. Biol.*, 149:381–424, 1991.

[33] R. J. De Boer, L. A. Segel, and A. S. Perelson. Pattern formation in one and two-dimensional shape-space. *J. Theor. Biol.*, 155:295–233, 1992.

[34] R. J. De Boer, J. D. van der Laan, and P. Hogeweg. In F. J. Varela and W. D. Stein, editors, *Thinking about Biology*, New-York, 1992. Addison-Wesley.

[35] L. Dienes and E. W. Schoenheit. *Am. Rev. Tuberc.*, 20(92), 1929. (quoted in [125], page 55).

[36] P. Ehrlich. *Klin. Jahrb.*, 6(299), 1897. (quoted in [125], page 51).

[37] P. Ehrlich. Die wertbemessung des diphtherieheilserums. *Klin. Jahb.*, 60(299), 1897. [English translation in "The Collected Papers of Paul Ehrlich, Vol. 2, pp. 107-125]. (quoted in [125], page 65).

[38] P. Ehrlich. On immunity with special reference to cell life. *Roy. Soc. Proc.*, 66:424–448, 1900.

[39] P. Ehrlich and J. Morgenroth. *Berl. Klin. Wochenschr.*, 36(481), 1899. (quoted in [125], page 260).

[40] P. Ehrlich and J. Morgenroth. *Berl. Klin. Wochenschr.*, 37(453), 1900. (quoted in [125], page 265).

[41] P. Ehrlich and J. Morgenroth. *Berl. Klin. Wochenschr.*, 38(251), 1901. (quoted in [125], page 264).

[42] H. N. Eisen and F. Karush. The interaction of purified antibody with homologous hapten. antibody valence and binding constant. *J. Am. Chem. Soc.*, 71:363–372, 1949.

[43] H. N. Eisen and G. W. Siskind. Variations in affinities of antibodies during the immune response. *Biochemistry*, 3, 1964.

[44] P. W. Erwin and F. Aladjem. The heterogeneity of antibodies with respect to equilibrium constants. calculation by a new method using delta functions, and analysis of the results. *Immunochemistry*, 13:873–883, 1976.

[45] J. D. Farmer, N. H. Packard, and A. S. Perelson. The immune system adaptation and machine learning. *Physica 22d*, pages 187–204, 1986.

[46] G. Fracastoro. *De Contagione et Contagiosis Morbis et Eorum Curatione*, pages 60–63. Putnam, New York, 1930. W. C. Wright trans. (quoted in [125], page 8).

[47] P. G. H. Gell and A. Kelus. *Nature (London)*, 201(687), 1964. (quoted in [125], page 254).

[48] E. A. Goidl, W. E. Paul, and B. Benacerraf. The effect of antigen dose and time after immunization on the amount and affinity of anti-hapten antibody. *J. Immunol.*, 100:371–381, 1968.

[49] E. S. Golub and D. R. Green. *Immunology: a synthesis*, chapter 1, page 3. Sinauer Associates, Inc., 1991.

[50] E. S. Golub and D. R. Green. *Immunology: a synthesis*, chapter 4, page 59. Sinauer Associates, Inc., 1991.

[51] N. Gunther and G. W. Hoffmann. Qualitative dynamics of a network model of regulation of the immune system: a rationale for the IgM to IgG switch. *Journal of Theoretical Biology*, 84:815–845, 1984.

[52] T. N. Harris, E. Grimm, E. Mertens, and W. E. Ehrich. *J. Exp. Med.*, 81(73), 1945. (quoted in [125], page 56).

[53] F. Haurowitz. *Nature (London)*, 205(847), 1930. (quoted in [125], page 69).

[54] M. Heidelberg and F. E. Kendall. *Science*, 72(252), 1930. (quoted in [125], page 64).

[55] E. Hertzfeld and R. Klinger. *Biochem. Z.*, 85(1), 1918. (quoted in [125], page 62).

[56] G. W. Hoffmann. A theory of regulation of self-nonself discrimination in an immune network. *European Journal of Immunology*, 5:638–647, 1975.

[57] G. W. Hoffmann. A mathematical model of the stable states of a network theory of self-regulation. *Lecture notes in biomathematics*, 32:239–257, 1979.

[58] G. W. Hoffmann. On network theory and H-2 restriction. In N. Warner, editor, *Contemporary topics in immunobiology, vol 11*, pages 185–226, 1980.

[59] G. W. Hoffmann. The application of stability criteria in evaluating network regulation models. In Charles De Lisi and Jean Hiernaux, editors, *Regulation of Immune Response Dynamics*, pages 137–162. CRC Press, 1982.

[60] G. W. Hoffmann, T. A. Kion, R. B. Forsyth, K. G. Soga, and A. Cooper-Willis. The N-dimensional network. In A. S. Perelson, editor, *Theoretical immunology, Part Two*, pages 291–319. Addison-Wesley, Redwood City, CA., 1988.

[61] G. W. Hoffmann, M. J. Lyons, D. J. Mathewson, M. J. Paetkau, and G. K. Soga. Modeling the immune system network in two dimensions, N dimensions and one dimension. In E. P. Hofer, editor, *Proceedings of the international seminar on modeling and signal processing of cellular systems*, 1992.

[62] G. W. Hoffmann and F. Tufaro. Serological distance coefficients. *Immunology letters*, 22:83–90, 1989.

[63] S. B. Hooker and W. C. Boyd. *J. Immunol.*, 21(113), 1931. (quoted in [125], page 65).

[64] S. B. Hooker and W.C. Boyd. *J. Immunol.*, 23(465), 1932. (quoted in [125], page 65).

[65] C. S. Hsu. Estimation of antibody affinity via reduced order modeling. In A. S. Perelson, editor, *Theoretical immunology, Part Two*, pages 265–289. Addison-Wesley, Redwood City, CA., 1988.

[66] N. K. Jerne. *Proc. Natl. Acad. Sci. USA*, 41(849), 1955. (quoted in [125], page 77).

[67] N. K. Jerne. The immune system. *Sci. Amer.*, pages 52–60, July 1973.

[68] N. K. Jerne. Towards a network theory of the immune system. *Ann. Immunol.*, 125 C:373–389, 1974. (Inst. Pasteur.).

[69] N. K. Jerne. The immune system: a web of V domains. The Harvey Lectures, 20th march 1975.

[70] N. K. Jerne. Idiotypic networks and other preconceived ideas. *Immun. Rev.*, 79(5), 1984.

[71] T. D. Jones and J. R. Mote. *N. Engl. J. Med.*, 210(120), 1934. (quoted in [125], page 56).

[72] P. Jordan. *Z. Immunitaetsforsch.*, 97(330), 1940. (quoted in [125], page 76).

[73] F. Karush. *Trans. N.Y. Acad. Sci.*, 20(581), 1958. (quoted in [125], page 70).

[74] F. Karush. Immunologic specificity and molecular structure. *Advanc. Immunol.*, 2:1–40, 1962.

[75] M. Kaufman and R. Thomas. Model analysis of the bases of multistationarity in the humoral immune response. *J. Theor. Biol.*, 129:141–162, 1987.

[76] M. Kaufman, J. Urbain, and R. Thomas. Towards a logical analysis of the immune response. *J. Theor. Biol.*, 114:527–561, 1985.

[77] A. Knorr. *Muench. Med. Wochenschr.*, 45(321,362), 1898. (quoted in [125], page 62).

[78] R. Koch. *Untersuchungen ueber die Aetiologie der Wundinfectionskrankheiten.* Vogel, Leipzig, 1878. (quoted in [125], page 41).

[79] H. Kossel. *Berl. Klin. Wochenschr.*, 35(152), 1898. (quoted in [125], page 262).

[80] R. Kraus. *Wien. Klin. Wochenschr.*, 10(736), 1897. (quoted in [125], page 51).

[81] K. Landsteiner and M. Reich. *Centralbl. Bakteriol.*, 39(712), 1905. (quoted in [125], page 67).

[82] J. Lederberg. *Science*, 129(1649), 1959. (quoted in [125], page 79).

[83] S. Lipschutz. *General Topology.* Shaum's Outline Series. McGraw-Hill, New-York, 1965. (quoted in [62].

[84] A. L. Locke, E. R. Main, and E. F. Hirsch. *Arch. Pathol.*, 2(437), 1926. (quoted in [125], page 63).

[85] W. H. Manwaring. *J. Immunol.*, 12(177), 1926. (quoted in [125], page 63).

[86] W. H. Manwaring. *J. Immunol.*, 19(155), 1930. (quoted in [125], page 64).

[87] W. H. Manwaring. *Science*, 72(23), 1930. (quoted in [125], page 64).

[88] C. Mather. The Angel of Bethesda. Am. Antiquarian Soc., Barre, Massachussets, 1924. (quoted in [125], page 13).

[89] D. Mathewson. Master's thesis, University of British Columbia.

[90] D. Mathewson and G. Hoffmann. Manuscript in preparation.

[91] Donald Mathewson. (Personal communication).

[92] H. Mercurialis. *De Morbis Puerorum*, volume Book I, pages 17–21. Basel, 1584. (quoted in [125], page 9).

[93] E. Metchnikoff. *Virchows Arch.*, 96(177), 1884. (quoted in [125], page 41).

[94] S. Mudd. *J. Immunol.*, 23(423), 1932. (quoted in [125], page 69).

[95] A. U. Neumann and G. Weisbuch. Dynamics and topology of idiotypic networks. *Bull. of Math. Biol.*, 54(5):699–726, 1992.

[96] A. U. Neumann and G. Weisbuch. Window automata analysis of population dynamics in the immune system. *Bull. of Math. Biol.*, 54(1):21–44, 1992.

[97] A. Nisonoff and D. Pressman. Heterogeneity and average combining constants of antibodies from individual rabbits. *J. Immunol.*, 80:417–428, 1958.

[98] G. Nuttall. *Z. Hyg.*, 4(353), 1888. (quoted in [125], page 47).

[99] I. I. Ostromuislensky. *J. Russ. Phys. Chem. Soc.*, 47(263), 1915. (quoted in [125], page 68).

[100] I. I. Ostromuislensky and D. I. Petrov. *J. Russ. Phys. Chem. Soc.*, 47(301), 1915. (quoted in [125], page 68).

[101] J. Oudin and M. Michel. *C. R. Hebd. Seances Acad. Sci.*, 257(805), 1963. (quotes in [125], page 254).

[102] G. Parisi. Immunological memory in a network perspective. In R. Livi, S. Ruffo, S. Ciliberto, and M. Buiatti, editors, *Chaos and complexity*, pages 394–406. World Scientific, 1989.

[103] L. Pasteur, C. Chamberland, and E. Roux. *C. R. Hebd. Seances Acad. Sci.*, 90(239), 1880. (quoted in [125], page 43).

[104] L. Pasteur, J. Joubert, and C. Chamberland. *C. R. Hebd. Seances Acad. Sci.*, 86(1037), 1878. (quoted in [125], page 41).

[105] L. Pauling. *Science*, 92(77), 1940. (quoted in [125], page 76).

[106] L. Pauling. A theory of the structure and process of formation of antibodies. *J. Am. Chem. Soc.*, 62(2643), 1940.

[107] L. Pauling, D. Pressman, and A. L. Grossberg. The serological properties of simple substances. vii. a quantitative theory of the inhibition by haptens of the precipitation of heterogeneous antisera with antigens and comparison with experimental results for poly-haptenic simple substances and for azoproteins. *J. Am. Chem. Soc.*, 66:784–795, 1944.

[108] J. Percus. Polydiversity in immune networks. In A. S. Perelson, editor, *Theoretical immunology, Part Two*, volume III of *SFI studies in the science of complexity*, pages 345–358. Addison-Wesley, Redwood City, CA., 1988.

[109] A. S. Perelson. Toward a realistic model of the immune response. In A. S. Perelson, editor, *Theoretical immunology, Part Two.* Addison-Wesley, Redwood City, CA., 1988.

[110] A. S. Perelson and G. F. Oster. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of Theoretical Biology*, 81:645–670, 1979.

[111] R. Pfeiffer. *Z. Hyg.*, 18(1), 1895. (quoted in [125], page 48).

[112] R. Pfeiffer and E. Friedberger. *Berl. Klin. Wochenschr.*, 39(4), 1902. (quoted in[125], page 267).

[113] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in Fortran*, chapter 16, pages 704–715. Cambridge University Press, 1992.

[114] G. Ramon. *R. Seances Soc. Biol. Ses Fil.*, 102(287, 379, 381), 1929. (quoted in [125], page 63).

[115] Rhazes. A treatise on the small-pox and measles. (W. A. Greenhill, trans.), London, 1848. (quoted in [125], page 6).

[116] A. R. Rich and M. R. Lewis. *Bull. John Hopskins Hosp.*, 50(115), 1932. (quoted in [125], page 56).

[117] P. H. Richter. A network theory of the immune response. *European Journal of Immunology*, 5:350–354, 1975.

[118] P. H. Richter. The network idea and the immune response. In A. S. Perelson G. I. Bell and G. H. Pimbley, editors, *Theoretical Immunology*, pages 539–569. Marcel Dekker, New York, 1978.

[119] P. H. Richter. The network idea and the immune response. In G. I. Bell, A. S. Perelson, and G. H. Pimbley, editors, *Theoretical Immunology*, pages 539–569. Dekker, 1978.

[120] I. Roitt, J. Brostoff, and D. Male. *Immunology*, chapter 7, page 7.6. Gower Medical Publishing, London, 2nd edition, 1989.

[121] E. Roux and L. Vaillard. *Ann. Inst. Pasteur, Paris*, 7(65), 1893. (quoted in [125], page 62).

[122] H. Sachs. *Ergeb. Allg. Pathol. Anat.*, 7(714, 768), 1901. (quoted in [125], page 268).

[123] R. S. Schweet and R. D. Owen. *J. Cell. Comp. Physiol.*, 50, Suppl. 1(199), 1957. (quoted in [125], page 74).

[124] L. A. Segel and A. S. Perelson. Computations in shape-space: a new approach to immune network theory. In A. S. Perelson, editor, *Theoretical immunology, Part Two.* Addison-Wesley, Redwood City, CA., 1988.

[125] A. M. Silverstein. *A history of immunoloy.* Academic Press, Inc., San Diego, California, 1989.

[126] A. M. Silverstein. *A history of immunoloy*, chapter 1, page 7. Academic Press, Inc., San Diego, California, 1989.

[127] A. M. Silverstein. *A history of immunoloy*, chapter 1, pages 10–11. Academic Press, Inc., San Diego, California, 1989.

[128] A. M. Silverstein. *A history of immunoloy*, chapter 1, pages 13–21. Academic Press, Inc., San Diego, California, 1989.

[129] A. M. Silverstein. *A history of immunoloy*, chapter 10, page 256. Academic Press, Inc., San Diego, California, 1989.

[130] F. A. Simon and F. M. Rackemann. *J. Allergy*, 439(451), 1934. (quoted in [125], page 56).

[131] D. Stauffer. Monte-Carlo simulation of Ising-like immunological shape-space. *Int. J. Mod. Phys. C.* (preprint).

[132] D. Stauffer and M. Sahimi. High-dimensional simulation of simple immunological models. (preprint).

[133] D. Stauffer and M. Sahimi. High-dimensional and very large cellular automata for immunological shape space. *Int. J. Mod. Phys. C*, 4:401–408, 1993.

[134] D. Stauffer and G. Weisbuch. *Physica A*, 180:42–52, 1992.

[135] J. Stewart and F. Varela. Morphogenesis in shape-space: Elementary meta-dynamics in a model of the immune network. *Journal of Theoretical Biology*, 153:477–498, 1991.

[136] D. W. Talmage. *Annu. Rev. Med.*, 8(239), 1957. (quoted in [125], page 78).

[137] D. W. Talmage. *Science*, 129(1643), 1959. (quoted in [125], page 79).

[138] F. H. Thiele and D. Embleton. *Z. Immunitaetsforsch.*, 20(1), 1814. (quoted in [125], page 68).

[139] Thucydides. *The Peloponnesian War.* Modern Library, New York, 1934. Crawley, trans. (quoted in [125], page 2).

[140] W. W. C. Topley. *J. Pathol. Bacteriol.*, 33(339), 1930. (quoted in [125], page 64).

[141] W. W. C. Topley. *J. Pathol. Bacteriol.*, 33(341), 1930. (quoted in [125], page 69).

[142] F. J. Varela, A. Coutinho, B. Dupire, and N. M. Vag. Cognitive networks: immune, neural and otherwise. In A. S. Perelson, editor, *Theoretical immunology, Part Two*, pages 359–375. Addison-Wesley, Redwood City, CA., 1988.

[143] R. Virchow. *Cellular Pathology.* Dover, New-York, 1971. (quoted in [125],page 40).

[144] E. von Behring and S. Kitasato. *Dtsch. Med. Wochenschr.*, 16(1113), 1890. (quoted in [125], page 49).

[145] E. von Behring and E. Wernicke. *Z. Hyg.*, 12(10, 45), 1892. (quoted in [125], page 49).

[146] A. von Wasserman. *Z. Hyg. Infektionskr.*, 37(173), 1901. (quoted in [125], page 268).

[147] R. G. Weinand. Somatic mutation, affinity maturation and the antibody repertoire: a computer model. *J. Theor. Biol.*, 143:343–382, 1990.

[148] R. G. Weinand. Somatic mutation and the antibody repertoire: a computational model of shape-space. In A. S. Perelson and S. A. Kauffman, editors, *Molecular evolution on rugged landscapes*, volume IX of *SFI studies in the science of complexity*, pages 215–236. Addison-Wesley, Redwood City, CA, 1991.

[149] G. Weisbuch. Dynamical behaviour of discrete models of jerne's network. In H. Atlan and I. R. Cohen, editors, *Theories of immune networks*, volume 46 of *Springer Series in Synergetics*, Heidelberg, 1989. Springer-Verlag Berlin.

[150] G. Weisbuch. A shape space approach to the dynamics of the immune system. *J. Theor. Biol.*, 143:507–522, 1990.

[151] G. Weisbuch, R. J. De Boer, and A. S. Perelson. Localized memories in idiotypic networks. *J. Theor. Biol.*, 146:483–499, 1990.

[152] T. P. Werblin and G. W. Siskind. Distribution of antibody affinities: technique of measurement. *Immunochemistry*, 9:987–1011, 1972.

[153] A. E. Wright and S. R. Douglas. *Proc. R. Soc. London, Ser. B*, 72(364), 1903. (quoted in [125], page 53).

[154] E. Yee. Reconstruction of the antibody affinity distribution from experimental binding data by a minimum cross-entropy procedure. *Journal of Theoretical Biology*, 153:205–227, 1991.

[155] H. Zinsser. *J. Exp. Med.*, 34(495), 1921. (quoted in [125], page 55).

[156] H. Zinsser. *J. Exp. Med.*, 41(159), 1925. (quoted in [125], page 55).