

# Coarse graining in genetic dynamics: A renormalization group analysis of a simple genetic system

C.R. Stephens\*, C. Chryssomalakos\*\*, and Adolfo Zamora\*\*\*

*Instituto de Ciencias Nucleares, U.N.A.M.,*

*Circuito Exterior, Apartado Postal 70-543, México D.F. 04510, México.*

\*stephens@nuclecu.unam.mx, \*\*chryss@nuclecu.unam.mx, \*\*\*zamora@nuclecu.unam.mx

Recibido el 19 de septiembre de 2003; aceptado el 15 de abril de 2004

We show how the idea of coarse graining can be applied fruitfully to the area of genetic dynamics, both in the context of “effective” theories - leading to more appropriate effective degrees of freedom with which to describe the dynamics - as well as in terms of integrating out degrees of freedom, using the Renormalization Group as a systematic calculational scheme. Specializing to dynamics in the presence of selection and mutation we show how the Renormalization Group can be implemented at the level of a transfer matrix-type description. Further, we present an explicit exact RG transformation for the simple case of a one gene-two allele system, solving for its fixed points and the asymptotic behaviour of the system in the vicinity of these fixed points. Renormalization Group, Genetic Dynamics, Coarse graining, Selection, Mutation, Evolution.

Se muestra como la idea de “coarse graining” puede ser aplicado exitosamente al area de la dinámica genética, tanto en el contexto de teorías “efectivas” - llevando a grados de libertad mas adecuados con que describir la dinámica - como en términos de la integración de grados de libertad, usando el grupo de renormalización como un esquema sistemático computacional. Particularizando la dinámica en la presencia de mutación y selección mostrando como el grupo de renormalización puede ser implementado al nivel de una descripción de tipo matriz de transferencia. Adicionalmente, se presenta una transformación del grupo de renormalización exacta y explícita para el caso de un sistema de un gen y dos alelos, solucionandola para obtener los puntos fijos correspondientes y el comportamiento asintótico del sistema cerca a estos puntos fijos.

*Descriptores:* Grupo de renormalización, dinámica genética, selección, mutación, evolución.

PACS: 05.10.Cc, 87.10.+e, 87.23.Kg, 89.75.-k

## 1. Introduction

The Renormalization Group (RG) is arguably the most powerful tool developed to date for analysing, both qualitatively and quantitatively, systems with many degrees of freedom. Its value is manifest in the large spectrum of successful applications associated with it, ranging from relativistic quantum field theory to the asymptotics of differential equations. Recently, it has also been applied to the field of genetic dynamics [1]. With a mainly physics audience in mind, in this article we will further consider its application in this area. In particular, we will present an explicit exact RG transformation and solve the associated RG equations in the context of a simple genetic model.

By genetic dynamics we mean the dynamics of populations of string-, or tree-like objects whose evolution is governed by a set of genetic operators. The most common operators, which one may think of as caricatures of the corresponding operators found in real biological systems, are: selection, mutation and recombination. Selection and mutation have been studied by physicists (see for example [2]). Recombination however remains relatively untouched, although it has been intensively studied in biology (see for example [3]). The basics of genetic dynamics are relevant to different fields, such as population genetics, and associated fields, and evolutionary computation. However, the underlying nature of the systems being modelled in the different areas can be radically different. A common thread on the other hand is that one is in practice dealing with many, many degrees of freedom and hence the normal motivation for applying coarse graining methods, of reducing degrees of freedom, is valid.

For instance, a typical protein has  $O(10^4)$  aminoacids. As there are 21 aminoacids the number of associated states is  $O(21^{10^4})$ .

Besides its calculational utility, coarse graining can also be of great utility in terms of identifying, qualitatively, the collective, or “effective”, degrees of freedom of a system which consists of very many “microscopic” degrees of freedom. This may often times help in seeking a more appropriate “effective” theory with which to do calculations. Hydrodynamics would be a case in point. It may also lead to conceptual insights of great importance. One only need remember that application of the RG in the context of critical phenomena leads to very deep insights, and “explains” scaling and universality, without having to calculate explicitly any number.

The structure of the paper will be as follows: in Sec. 2 we will give a brief introduction to the topic of genetic dynamics, discussing some basic features. In Sec. 3 we give an overview of coarse graining in the context of genetic dynamics showing how different coarse grainings are naturally associated with the different genetic operators. In Sec. 4 we introduce, in a formal context, the RG as a framework within which to understand the different coarse grainings introduced. In Sec. 5 we restrict attention to genetic dynamics with mutation and selection only, putting the dynamics in the context of a transfer matrix formalism. In Sec. 6, to illustrate the techniques and concepts developed and discussed, we apply the RG to a simple genetic system consisting of one gene evolving under the effect of mutation and selection. Finally, in Sec. 7 we draw some conclusions.

## 2. Basics of Genetic Dynamics

At its most basic level a genetic dynamics is a stochastic process that takes as input a population of “objects” (strings, trees etc.) - the “genotypes” - and a fitness function, at a given time, and gives as output the population at a later time. The objects live on a configuration space  $X$ , of dimensionality  $N_X$ , with elements  $i \in X$ . We denote a population by  $\mathbf{P} = (n_1, n_2, \dots, n_{N_X})$ , where  $n_i$  represents the proportion of objects of type  $i$  in the population. Each object is assigned a quality or fitness via a fitness function  $f_X : X \rightarrow R^+$ . Population flows take place on  $X$ . One may also consider the dynamics in the space of populations rather than on  $X$ , the former being a simplex whose vertices correspond to the totally uniform population states, of which there are  $2^N$  for binary strings, while the simplex center corresponds to the completely random state where all strings are represented equally.

A dynamics is imposed via an evolution operator generated by the genetic operators and which can be written formally as

$$\mathbf{P}(t + 1) = \mathcal{T}(\mathbf{P}(t), \mathbf{p}) \mathcal{S}(\mathbf{P}(t), \mathbf{f}) \quad (1)$$

where  $\mathcal{T}(\mathbf{P}(t), \mathbf{p})$  is a matrix and  $\mathcal{S}(\mathbf{P}(t), \mathbf{f})$  is a vector. The transmission term  $\mathcal{T}(\mathbf{P}(t), \mathbf{p})$  describes the probability of transforming one object into another by mutation, crossover, or other genetic operators, the explicit transmission mechanism being encoded by the parameters  $\mathbf{p}$ . The term  $\mathcal{S}(\mathbf{P}, \mathbf{f})$  describes the selection forces acting on  $\mathbf{P}$  with the parameters  $\mathbf{f}$  determining the fitness function. For an infinite population (1) describes the evolution of the probability distribution while, for a finite system, it describes the relationship between an actual population state at time  $t$  and the expected population state at time  $t + 1$ .

In this paper, for simplicity, we will restrict attention to the evolution of strings (“chromosomes/proteins”) of fixed length, consisting of  $N$  bits (“genes”). Also, for simplicity, we will assume that the bits take only binary values (“alleles”), though nothing we shall present depends on this fact. In this case, the most natural representation of  $X$  is a  $N$ -dimensional hypercube, where the natural metric is Hamming distance (The Hamming distance between two strings is the number of bits that differ between the strings. For example, the Hamming distance between 100 and 010 is two.), strings associated with adjacent vertices being Hamming distance one apart. The  $N$  string loci are taken to be a complete orthonormal basis for the hypercube.

To make (1) more explicit one needs to specify the particular “genetic” operators that generate the dynamics. As mentioned we will restrict attention here to the three canonical operators - selection, mutation and recombination. In this case the evolution operator depends on the reproductive fitness landscape,  $\mathbf{f}$ , the population  $\mathbf{P}$  and the set of parameters,  $\mathbf{p}$ , that govern the other genetic operators; e.g. mutation and recombination probabilities. For selection,  $P_i(t+1) \equiv P'_i(t) = F_{ij} P_j(t)$ , where  $F_{ij}$  is the fitness matrix and  $P_i(t)$  is the probability of finding the string  $i$  at time  $t$ . A typical selection scheme is proportional selection. In this

case the fitness matrix is  $F_{ij} = (f_i/\bar{f}(t))\delta_{ij}$ , where  $\bar{f}(t)$  is the average population fitness. It is usually considered as a unary operator.

Mutation, typically, is such that every string bit flips to its complement with probability  $p$  every generation and is also a unary operator. Recombination, however, is almost always a binary operator (although higher cardinality can be considered). Recombination is implemented by taking a certain number of bits from one “parent” string and the complement from another “parent” string to form a “child” string. For example, one can form 1111 from parents 1010 and 0101 by taking the first and third bits from 1010 and the complementary second and fourth bits from 0101. The bits taken from the different parents can be specified using a recombination “mask”,  $m$ . For instance, in the above example the recombination mask is 1010 which signifies take bits one and three from the first parent (specified by the position of the ones) and two and four from the second.

The resultant dynamical equation describing the evolution of the probability distribution for this system is

$$P_i(t + 1) = \sum_j \mathcal{W}_{ij} P_j^c(t) \quad (2)$$

where  $P_i^c(t)$  is the probability to find strings of type  $i$  after selection and crossover. The mutation matrix,  $\mathcal{W}$ , has matrix elements  $\mathcal{W}_{ij} = p^{d_{ij}^H} (1-p)^{N-d_{ij}^H}$ , where  $d_{ij}^H$  is the Hamming distance between the two strings. As in the case of the more abstract (1), (2) also applies for a finite population if we interpret the left hand side of (2) as the expected proportion of genotype  $i$  to be found at  $t + 1$ , while any  $P_i(t)$  on the right hand side are to be considered as the actual proportions found at  $t$ .

In order to have a closed set of equations we need to specify  $P_i^c(t)$  in terms of the underlying  $P_i(t)$ . The relation is

$$P_i^c(t) = (1-p_c)P_i(t) + \sum_m \sum_j \sum_k \lambda_{ijk}(m) P_j'(t) P_k'(t) \quad (3)$$

where  $p_c$  is the probability to implement recombination and  $P_i'(t)$  is the probability to select  $i$ .  $P_i' = (f_i/\bar{f}(t))P_i$  for proportional selection.  $\lambda_{ijk}(m)$  is an interaction term between strings, that depends on the particular crossover mask  $m$ , and  $\sum_{m=1}^{2^N}$  is the sum over all possible recombination masks. Note that the interaction constants,  $\lambda_{ijk}(m)$ , are independent of the population. For a given target string,  $i$ ,  $\lambda_{ijk}(m)$  is a  $2^N$ -dimensional matrix. The matrix is very sparse however, having only  $O(2^N)$  non-zero elements for a given target string. Thus, this “microscopic” representation in terms of individual string states is very inefficient, there being very few ways of creating a given target by recombination of strings. The vast majority of string recombination events are neutral in that they lead to no non-trivial interaction.

The sum over masks takes into account the different ways genetic material can be combined from the different parent strings to form a given child string. As an interaction,  $\lambda_{ijk}$  takes into account different interaction events that can result in the gain of a given string. Taking as target the string 111, for example, recombining 110 and 001 with a mask 110 leads

to the formation of 111. It is not difficult to write down explicit forms for the interaction constants - see for example [1].

Unlike mutation, for recombination Hamming distance is not a natural metric. For example, consider two parent strings 111111111 and 000000000. A one-point crossover implemented between the last two bits leads to offspring 111111110 and 000000001, which are Hamming distance one from the respective parents. An equally probable crossover between the fifth and six bits, however, leads to 111110000 and 000011111, which are Hamming distance five away from the parents.

The Eqs. (2) and (3) yield an exact expression for the probability distribution governing the evolution for arbitrary selection, mutation and crossover. It takes into account exactly the effects of destruction and construction of strings.

### 3. Coarse-Graining and Coordinate Transformations

We have shown that the dynamics of our string system is described by  $2^N$  coupled, non-linear difference equations representing the microscopic degrees of freedom, i.e. the strings themselves. This is an exact representation. However, obtaining useful information from these equations is a highly non-trivial undertaking. The field of population genetics has been concerned precisely with this task for over 80 years now, and there most progress has been at the level of toy models with very simple fitness landscapes or a small number of genetic loci (string bits) - often at the level of only 2 or 3 bits.

There are two basic motivations for performing a coarse graining of a system with many degrees of freedom. One is: to write the dynamics in terms of more appropriate effective degrees of freedom, i.e. to write an “effective” dynamics, while the other is to solve for the dynamics by considering a related system with fewer degrees of freedom. In this sense we are using the word coarse graining in what would normally be considered to be two different contexts. For instance, writing the dynamics of a set of coupled harmonic oscillators in terms of normal modes would be an example of writing dynamics in terms of more appropriate degrees of freedom - the normal modes being collective (“coarse grained”) degrees of freedom composed from the underlying microscopic degrees of freedom - the oscillators themselves. Of course, the normal way of thinking of this exact coarse graining is as a symplectic transformation. Here, the word coarse graining is meant to indicate that it is a description in terms of collective degrees of freedom. The mapping between the two is exact. Another simple example would be that of describing the dynamics of a macroscopic metal sphere in terms of its position coordinates and Euler angles. Once again these degrees of freedom are collective degrees of freedom relative to the underlying more fundamental atomic degrees of freedom. The mapping however in this case is approximate as we forego information about all atomic degrees of freedom. Mathematically, this is described by a projection rather than a symplectic transformation as above.

It is this concept of a projection on the configuration space that is at the heart of the second meaning of coarse

graining - “integrating out” degrees of freedom - which is more directly associated with considerations of the RG. Once again, a coarse graining may be exact or approximate, though the number of explicit exact coarse graining transformations is very limited. An example would be the one-dimensional Ising model [4]. The motivation behind this type of coarse graining is that if you can map a system onto an analogous system with fewer degrees of freedom then the latter should be easier to solve, or at least approximately solve. Both types of coarse graining naturally appear in genetic dynamics as we shall see.

We will start off with coarse grainings associated with more natural collective degrees of freedom. A very simple, but relevant, example is that of phenotypic dynamics. Fitness, almost by definition, acts at the level of the phenotype (The observable manifestation of the genotype) hence if the dynamics is engendered by nothing other than pure selection it is natural to perform a coarse graining from genotype to phenotype. As a concrete example consider a fitness landscape where the fitness is given by the number of ones on the string (a simple paramagnet). In this case the dynamics can be rewritten in terms of the  $N$  phenotypes rather than  $2^N$  genotypes. The equation of motion for selection only is then

$$P_n(t+1) = \frac{n}{\bar{n}(t)} P_n(t) \quad (4)$$

where we denote phenotypes by  $n$ , the number of ones, and  $\bar{n}(t)$  is the average number of ones in the population at time  $t$ . The solution of these  $n$  difference equations is

$$P_n(t) = \frac{n^t P_n(0)}{\sum_{n=0}^N n^t P_n(0)} \quad (5)$$

Another example is that of the Eigen model [5], where the fitness landscape is degenerate for all genotypes except one, the master sequence. At the level of selection only, given that there are only two phenotypes, there is a reduction in the size of the configuration space from  $2^N$  to 2, i.e. a reduction in the number of degrees of freedom from  $N$  to 1. However, if we include in the effect of mutation we see there is an induced breaking of the genotype-phenotype symmetry due to the fact that strings close to the master sequence in Hamming distance have a higher “effective” fitness [6, 7]. In both these cases a natural coarse graining is suggested by the form of the fitness landscape, which itself is intimately related to the selection operator. Note that, although the genotype to phenotype mapping is a true coarse graining, in the sense that it is a projection on  $X$ , it is a very simple one as it is an exact coarse graining associated with a symmetry. Information is lost in the sense that after the projection we can no longer enquire as to the dynamical evolution of a particular genotype.

In the case of selection and mutation the dynamical equations are essentially linear, the apparently non-linear average fitness only acting as a normalization factor, and, as we will see, the resulting selection/mutation problem can be recast in the guise of a two-dimensional, inhomogeneous statistical mechanics problem, where powerful techniques such as the transfer matrix approach can be invoked. In this case the natural effective degrees of freedom are the normal modes of the

mutation-selection matrix, i.e. its eigenvectors. In the case of pure mutation a Walsh transformation (Fourier transform on binary numbers) yields the eigenvectors. In both cases the transformation is a coordinate transformation rather than a projection and no information is lost.

Explicitly, the *Walsh basis*  $\psi$ , consists of *Walsh functions*,

$$\psi_I(x) = 1/\sqrt{|X|} \prod_{j \in I} x_j,$$

where  $I$  is a subset of  $\{1, \dots, N\}$  and  $x_j = \pm 1$ . The Walsh functions are normed and orthogonal and of *order*  $|I|$ , the number of loci that are multiplied. The Walsh-transform,  $\hat{f}$ , of a function  $f$  is defined implicitly by

$$f(i) = \sum_I \hat{f}(I) \psi_I(i)$$

Multiplying with  $\psi_K(x)$  and summing over all  $i \in X$  we obtain

$$\begin{aligned} \sum_{i \in X} f(i) \psi_K(i) &= \sum_I \hat{f}(I) \sum_{i \in X} \psi_I(i) \psi_K(i) \\ &= \sum_I \hat{f}(I) \delta_{IK} = \hat{f}(K) \end{aligned} \tag{6}$$

In matrix form,  $\hat{f} = \Psi f$ , where the matrix  $\Psi$  has the Walsh functions  $\psi_K$  as its rows. As mentioned, the Walsh functions are eigenfunctions of the mutation operator  $\mathcal{W}$  that satisfies  $\mathcal{W}\psi_I = (1 - 2|I|/N)\psi_I$ . The mutation operator is therefore diagonal in the Walsh basis. Equation (2) reads in these coordinates

$$\hat{P}_I(t+1) = \sum_i (\Psi \mathcal{W})_{Ii} P_i^c = \left(1 - 2\frac{|I|}{N}\right) \hat{P}_I^c \tag{7}$$

For example, in the case of one gene

$$\mathcal{W}_1 \equiv \begin{pmatrix} (1-p) & p \\ p & (1-p) \end{pmatrix}$$

with  $\psi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  and  $\psi_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  one finds  $\mathcal{W}_1 \psi_0 = \psi_0$  and  $\mathcal{W}_1 \psi_1 = (1-2p)\psi_1$ . The Walsh basis will be particularly useful if the transformed selection-crossover term  $\hat{P}_I^c$  also has a simple form. The explicit Walsh matrix for 3 bits is seen below (Note that although we have defined above the Walsh functions in terms of an alphabet 1, -1 the Walsh matrix is the same for alphabet 1, 0 the only difference for the latter being how the Walsh function is defined.). Note that the last row corresponds to the constant eigenvector.

$$\Psi = \frac{1}{\sqrt{8}} \begin{array}{c|cccccccc} & 111 & 110 & 101 & 011 & 100 & 010 & 001 & 000 \\ \hline 111 & -1 & +1 & +1 & +1 & -1 & -1 & -1 & +1 \\ 110 & +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ 101 & +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ 011 & +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ 100 & -1 & -1 & -1 & +1 & -1 & +1 & +1 & +1 \\ 010 & -1 & -1 & +1 & -1 & +1 & -1 & +1 & +1 \\ 001 & -1 & +1 & -1 & -1 & +1 & +1 & -1 & +1 \\ 000 & +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \end{array} \tag{8}$$

A third, and much more non-trivial example, is associated with the recombination operator. The fact that the string representation for recombination is very inefficient, due to the sparsity of the interaction matrix, is an indication that strings are not the natural effective degrees of freedom for recombination. Insight into what are more appropriate effective degrees of freedom can be gleaned by considering a simple example: To form the string 111 with a recombination mask 100 one can join strings 111, 110, 101 and 100 with either 111 or 011. In other words, for the first parent the second and third bit values are unimportant and for the second the first bit value is unimportant. Thus, it is natural to coarse grain over those strings that give rise to the desired target for a given mask. Such coarse-grained variables are known as “schemata”, and are conceptually equivalent to, for instance, “block spins” in traditional statistical mechanics RG applications, except in this case there are eminently good reasons as to why they need not be “local”. The marginal probability,  $P_\alpha(t)$ , represents the probability of finding the schema  $\alpha$  at time  $t$ . A specific schema is determined by summing over

those bit positions that are not part of the schema. One may denote such a bit position by a \*. Thus, 11\* represents the two strings 111 and 110. The number of definite bits of the schema defines its order,  $N_2$ , while the number of bits between the outermost defining bits, including the latter, defines its length. Thus, \*11\*0\*\* has  $N_2 = 3$  and  $l = 5$ .

Generally, if one picks, arbitrarily, a vertex in  $X$ , associated with a string  $i$ , one may perform a linear coordinate transformation,  $\Lambda : X \rightarrow \tilde{X}$ , to a basis consisting of all schemata that contain  $i$ . For instance, for two bits  $X = \{11, 10, 01, 00\}$ . Selecting the string 11 as our “preferred” vertex, we have  $\tilde{X} = \{11, 1*, *1, **\}$ . The invertible matrix  $\Lambda$  is such that  $\Lambda_{\alpha i} = 1 \iff i \in \alpha$ . We denote the resulting basis the Building Block Basis (BBB) [1]. Given the arbitrariness of the choice of vertex there are in fact  $2^N$  equivalent BBBs each transformable to any other by a permutation. For 3 bits the explicit transformation matrix for the basis with preferred vertex 111 is

$$\Lambda = \begin{matrix} & \begin{matrix} 111 & 110 & 101 & 011 & 100 & 010 & 001 & 000 \end{matrix} \\ \begin{matrix} 111 \\ 11* \\ 1*1 \\ *11 \\ 1** \\ *1* \\ **1 \\ *** \end{matrix} & \begin{matrix} \begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix} \end{matrix} \end{matrix} \tag{9}$$

The BBB is not orthonormal. Note that the vertex  $i$  by construction is a fixed point of this transformation. Apart from the vertex  $i$ , the points in  $\tilde{X}$ , being schemata, correspond to higher dimensional objects in  $X$ . For instance, considering the two-dimensional case,  $1*$  and  $*1$  are one-planes in  $X$  while  $**$  is the whole space. In the BBB one may transform (3) to find

$$\tilde{P}_i^c(t+1) = (1 - p_c)\tilde{P}_i'(t) + \sum_{m=1}^{2^N} \sum_{\beta, \gamma} \tilde{\lambda}_{\alpha\beta\gamma}(m)\tilde{P}_j'(t)\tilde{P}_k'(t) \tag{10}$$

where  $\tilde{\lambda}_{\alpha\beta\gamma}(m) = \Lambda_{\alpha i} \lambda_{ijk} \Lambda_{\beta j}^{-1} \Lambda_{\gamma k}^{-1}$ .

The advantage of this new representation is that the properties and symmetries of crossover are much more transparent. For instance,  $\tilde{\lambda}_{\alpha\beta\gamma}(m)$  is such that, for a given mask, only interactions between BBs that construct the target string or schema are non-zero, *i.e.*,  $\tilde{\lambda}_{\alpha\beta\gamma}(m) = 0$ , unless  $\gamma$  corresponds to a schema which is the complement of  $\beta$  with respect to  $\alpha$ . Also,  $\tilde{\lambda}_{\alpha\beta\gamma}(m) = 0$  unless  $\beta$  is equivalent to  $m$ , where equivalent means that, for any 1 in the mask, we have a 1 at the corresponding locus in  $\beta$  and for any 0 we have a  $*$ . These two important properties mean that the summations over  $\beta$  and  $\gamma$  in (10) disappear to leave only the sum over masks with an “interaction” constant  $p_c(m)$  which depends only on the mask. For example, for two bits, if we choose as vertex 11, then 11 may interact only with  $**$ , while  $1*$  may interact only with  $*1$ .

In  $X$  this has the interesting interpretation that for a target schema,  $\alpha$ , of dimensionality  $(N - d)$ , only geometric objects “dual” in the  $d$ -dimensional subspace of  $X$  that corresponds to  $\alpha$  may interact. In other words, a  $k$ -dimensional object recombines only with a  $(N - d - k)$ -dimensional object. Additionally, a  $(N - d)$ -dimensional object may only be formed by the interaction of higher dimensional objects. In this sense interaction is via the geometric intersection of higher dimensional objects. For example, the point 11 can be formed by the intersection of the two lines  $1*$  and  $*1$ . Similarly, 1111 can be formed via intersection of the three-plane  $1***$  with the line  $*111$  or via the intersection of the two two-planes  $11**$  and  $**11$ .

As mentioned, one of the primary advantages of the BBB representation is that the sums over  $j$  and  $k$  in equation (3)

disappear thus obtaining

$$P_i^c(t) = (1 - p_c)P_i'(t) + \sum_{m=1}^{2^N} p_c(m)P_{i_m}'(t)P_{i_{\bar{m}}}'(t) \tag{11}$$

where  $P_{i_m}'(t)$  is the probability to select the BB  $i_m$  (note that the mask uniquely specifies which element,  $i_m$ , of the BBB to choose) and  $P_{i_{\bar{m}}}'(t)$  is the probability to select the BB  $i_{\bar{m}}$ , which is uniquely specified as the complement of  $i_m$  in  $i$ . Both  $i_m$  and  $i_{\bar{m}}$  are elements of the BBB associated with  $i$ . The above equation clearly shows that recombination is most naturally considered in terms of the BBB. In the standard basis there were of the order of  $2^{2N}$  elements of  $\lambda_{ijk}$  to be taken into account for a given  $i$ . In the BBB there is only one term. Of course, the coarse grained averages of  $i_m$  and  $i_{\bar{m}}$  contain  $2^N$  terms, still, the reduction in complication is enormous. Thus, crossover naturally introduces the idea of a coarse graining, the associated effective degrees of freedom being the BBs we have defined. This is an important point as it shows that evolution is acting in the presence of crossover most naturally at the level of populations, the BBs representing populations with a certain degree of “kinship” to the target object.

Inserting (11) in (2) we can try to solve for the dynamics. However, in order to do that we must know the time dependence of  $i_m$  and  $i_{\bar{m}}$ . Although the number of BB basis elements is  $2^N$  we may generalize and consider the evolution of an arbitrary schema,  $\alpha$ . To do this we need to sum with  $\sum_{i \in \alpha}$  on both sides of equation (2). This can simply be done to obtain [8–10] again the form (2), where this time the index  $\alpha$  runs only over the  $2^{N_2}$  elements of the schema partition and where again  $\mathcal{W}_{\alpha\beta} = p^{d_{\alpha\beta}^H} (1 - p)^{N - d_{\alpha\beta}^H}$ . In this case however  $d_{\alpha\beta}^H$  is the Hamming distance between the two schemata. For instance, for strings with three loci the schemata partition associated with the first and third loci is  $\{1*1, 1*0, 0*1, 0*0\}$ . In this case  $d_{12}^H = 1$  and  $d_{14}^H = 2$ .  $P_\alpha^c(t) = \sum_{i \in \alpha} P_i^c(t)$  is the probability of finding the schema  $\alpha$  after selection and recombination. Note the form invariance of the equation after coarse graining. To complete the transformation to schema dynamics we need the schema analog of (11). This also can be obtained by acting with  $\sum_{i \in \alpha}$  on both sides of the equation. One obtains

$$P_\alpha^c(t) = (1 - p_c N_\alpha)P_\alpha'(t) + \sum_{m \in M_r} p_c(m)P_{\alpha_m}'(t)P_{\alpha_{\bar{m}}}'(t) \tag{12}$$

where  $\alpha_m$  represents the part of the schema  $\alpha$  inherited from the first parent and  $\alpha_{\bar{m}}$  that part inherited from the second.

$N(\alpha)$  is the number of crossover masks that affect  $\alpha$ , relative to the total number of masks with  $p_c(m) \neq 0$ , the set of such masks being denoted by  $\mathcal{M}_r$ . Obviously, these quantities depend on the type of crossover implemented and on properties of the schema such as defining length. Note that the BBB naturally coarse grains here to the BBB appropriate for the schema  $\alpha$  as opposed to the string  $x$ .

Thus, we see that the evolution equation for schemata is form invariant, there being only a simple multiplicative renormalization of the recombination probability  $p_c$ . This form invariance, shown in [8, 9], demonstrates that BB schemata in general are a preferred set of coarse grained variables and, more particularly, the BBB is a preferred basis in the presence of recombination. It has also been shown [11] that schemata, more generally, are the *only* coarse graining that leads to invariance in the presence of mutation and recombination.

Considering again the structure of (11) and (12) we see that variables associated with a certain degree of coarse graining are related to BB “precursors” at an earlier time, which in their turn ... and so on. This hierarchical structure terminates at order-one BBs as these are unaffected by crossover. Thus, for example, the level one BB combinations of 111, *i.e.*, BBs that lead directly upon recombination to 111 are:  $11* : **1$ ,  $1*1 : *1*$  and  $1** : *11$ . The level two BBs are  $1**$ ,  $*1*$  and  $**1$ . Thus, a typical construction process is that BBs  $1**$  and  $*1*$  recombine at  $t = t_1$  to form the BB  $11*$  which at some later time  $t_2$  recombines with the BB  $**1$  to form the sequence 111.

So we see then that different genetic operators are typically associated with different natural sets of effective degrees of freedom. In these cases all the coarse graining has been done, in one form or another in “space”. It is also possible to coarse grain in “time”. We can think of the dynamics of a genetic system also in a “space-time” setting. Here, we imagine a description of the system on a finite two-dimensional space (rather than the evolution of a finite one-dimensional space (the string). This space is of “volume”  $N \times t$  and varies in size as the system evolves. Our genetic dynamics can now be thought of in terms of two-dimensional spin configurations. In this setting we can consider a coarse graining by integrating out “columns” of spins, thus effecting a reduction to a system of “height”  $t$  but “length”  $N' < N$ . It is clear however that one may also effect a coarse graining, as we will explicitly demonstrate shortly, by integrating out “rows” of spins, thus coarse graining to a system of length  $N$  but of height  $t' < t$ . Both these coarse grainings have natural interpretations. It is possible to also envision a more general coarse graining where blocks of spins other than rows or columns are integrated out.

### 4. Renormalization Group

In the previous section we saw that coarse grained variables arise very naturally in genetic dynamics and gave as examples the genotype-phenotype map and schemata. We can formalize these considerations by introducing a general coarse graining operator  $\mathcal{R}(\eta, \eta')$  which coarse grains from the variable  $\eta$  to the variable  $\eta'$ . In this case

$$\begin{aligned} \mathcal{R}(\eta', \eta)P(\eta, t) &= P(\eta', t) \\ \mathcal{R}(\eta'', \eta)P(\eta, t) &= P(\eta'', t) \end{aligned} \tag{13}$$

However, given that  $\mathcal{R}(\eta'', \eta')P(\eta', t) = P(\eta'', t)$  we deduce that

$$\mathcal{R}(\eta, \eta'') = \mathcal{R}(\eta, \eta')\mathcal{R}(\eta', \eta'') \tag{14}$$

*i.e.* the space of coarse grainings has a semi-group structure. Thus, we see that one can naturally introduce the RG into the study of genetic dynamics. The naturalness of a particular RG transformation will be to a large extent determined by how the dynamics looks under this coarse graining.

Considering (1), for the pdf of the dynamics, then, given that  $\mathcal{R}(\eta', \eta)P_\eta(t) = P_{\eta'}(t)$  the dynamics under a coarse graining is governed by  $\mathcal{R}(\eta, \eta')(\mathcal{T}(\mathbf{P}_\eta(t), \rho) \mathcal{S}(\mathbf{P}_\eta(t), \mathbf{f}))$ , where  $\mathcal{S}(\mathbf{P}_\eta(t), \mathbf{f})$  and  $\mathcal{T}(\mathbf{P}_\eta(t), \rho)$  are the dynamical operators associated with the variables  $\eta$ . If this can be written in the form  $\mathcal{T}(\mathbf{P}_{\eta'}(t), \rho') \mathcal{S}(\mathbf{P}_{\eta'}(t), \mathbf{f}')$  with suitable “renormalizations”,  $\mathbf{f} \rightarrow \mathbf{f}'$  and  $\rho \rightarrow \rho'$  of the model’s parameters, then the dynamics is form covariant or invariant under this coarse graining. Note that we are here considering a more general notion of invariance than the idea of “compatibility” [11] (see [12] for a discussion of the relationship between the two). In the case of selection only, the coarse graining transforms the fitness

$$f_\eta \rightarrow f_{\eta'} = \mathcal{R}(\eta, \eta')f_\eta = \sum_{\eta \in \eta'} f_\eta P_\eta(t) / \sum_{\eta \in \eta'} P_\eta(t). \tag{15}$$

An important point to note here is that, generically, a coarse graining gives rise to a *time dependent* coarse-grained fitness.

In the case of recombination note also that the coarse graining operator associated with the BBs satisfies

$$\mathcal{R}(\eta, \eta') = \mathcal{R}(\eta^s, \eta'^s)\mathcal{R}(\eta^c, \eta'^c) \tag{16}$$

where  $\mathcal{R}(\eta^s, \eta'^s)$  represents the action of the coarse graining on the BB  $\mathcal{S}$  while  $\mathcal{R}(\eta^c, \eta'^c)$  represents the action on the BB  $\mathcal{C}$ .

### 5. A Transfer Matrix approach to Mutation-Selection

In this section we will restrict attention to mutation-selection systems implementing a generic RG transformation at the level of the transfer matrix. We begin with the dynamical equation

$$\mathbf{P}(t+1) = \frac{1}{f(t)} \overline{\mathbf{W}}_s(t) \mathbf{P}(t) \tag{17}$$

where the selection-mutation matrix  $\overline{\mathbf{W}}_s(t) = \overline{\mathbf{W}}\mathbf{F}(t)$ , noting an invariance of the dynamics under the transformation on the fitness landscape  $f_i \rightarrow Cf_i$ , where  $C$  is any positive constant. This equation looks superficially non-linear due to the presence of the factor  $f(t)$ . However, this factor

can be eliminated without loss of generality, and then reinstated, by passing to unnormalized variables,  $y_i(t)$ , such that  $P_i(t) = y_i(t) / \sum_i y_i(t)$ , that satisfy

$$\mathbf{y}(t+1) = \overline{\mathbf{W}}_s(t)\mathbf{y}(t) \tag{18}$$

The solution to this equation is

$$\mathbf{y}(t) = \overline{\mathbf{W}}_s^t(t)\mathbf{y}(0) \tag{19}$$

and the resultant solution for  $\mathbf{P}(t)$  is

$$\mathbf{P}(t) = \frac{\overline{\mathbf{W}}_s^t(t)\mathbf{P}(0)}{\sum_i W_{ij}^t P_j(0)} \tag{20}$$

Thus, we see that knowledge of the matrix  $\overline{\mathbf{W}}_s$  is sufficient to determine the dynamics, in particular its eigenvalues and eigenvectors. The solution in terms of the eigenvalues,  $\lambda_i$ , and eigenvectors,  $\mathbf{u}_i$ , of  $\overline{\mathbf{W}}_s$  is

$$\mathbf{P}(t) = \frac{a_1 \mathbf{u}_1 \lambda_1^t + a_2 \mathbf{u}_2 \lambda_2^t + \dots}{\sum_{\alpha} u_{1\alpha} + (a_2/a_1)(\lambda_2/\lambda_1)^t u_{2\alpha} + \dots} \tag{21}$$

The matrix  $\overline{\mathbf{W}}_s$  is equivalent to the transfer matrix in statistical mechanics and, in fact, may be put into precisely that form [13], where the Hamiltonian of the resulting inhomogeneous two-dimensional Ising system is

$$H(T) = - \sum_{i=0}^{t-1} \sum_{j=1}^N (J_y \sigma_j^{i+1} \sigma_j^i + T \ln f_i) - \frac{TNt}{2} \ln(p(1-p)) \tag{22}$$

where the ‘‘temperature’’ of the system is

$$1/T = \ln(p/(1-p))$$

and  $\ln f_i$  is the energy associated with a row of spins. One can also make a correspondence between other observables of interest and their thermodynamic analogs. For instance, relaxation into the stationary state is governed by an analog of the correlation length and given by

$$\tau^{-1} = \ln(\lambda_1/\lambda_2) \tag{23}$$

We may formally consider a coarse graining of the above system in either direction, space or time, in the genetic system. Considering the time direction we coarse grain the dynamics by integrating over certain rows of spins. In terms of the transfer matrix we consider a coarse-grained matrix  $W' = W^n$  implying that we are coarse graining over  $n$  time steps. Whether the system is naturally ‘‘renormalizable’’ or not depends on whether  $W'$  can be naturally written in terms of reparametrized versions of the parameters that form  $W$ . In general, for a mutation-selection system there are  $2^N$  parameters associated with the fitness landscape and  $2N$  possible mutation rates, if we accept the possibility of asymmetric mutation rates (*i.e.* for two strings  $i$  and  $j$  the probability that  $i$  mutates to  $j$  is not the same that  $j$  mutates to  $i$ ).

The transfer matrix  $\overline{\mathbf{W}}_s(t)$  evolves the system from  $t$  to  $t + 1$  or, put another way, links spin rows  $t$  and  $t + 1$ . Spin rows  $t$  and  $t + 2$  are linked by the matrix  $\overline{\mathbf{W}}_s'(t) = \overline{\mathbf{W}}_s(t+1)\overline{\mathbf{W}}_s(t)$ , the matrix  $\overline{\mathbf{W}}_s'(t)$  evolving the system over two time steps. More generally,  $\overline{\mathbf{W}}_s'(t, t')$  will be taken as the matrix that evolves the system from  $t'$  to  $t$ . These matrices satisfy

$$\overline{\mathbf{W}}_s'(t, t'') = \overline{\mathbf{W}}_s'(t, t')\overline{\mathbf{W}}_s'(t', t'') \tag{24}$$

and hence give a representation of the RG. Thus, we see how integrating out rows of spins at the level of the transfer matrix leads to a simple realization of the RG. In the next section we will consider a concrete realization of this mapping.

### 6. Exact RG for a One Gene-Two Allele System

We illustrate the RG introduced in the previous section with a concrete, simple example - that of a genetic system consisting of one gene (bit) with two alleles (values) with fitness values  $f_1$  and  $f_0$ . In this case the transfer matrix is

$$\begin{vmatrix} (1-p)f_1 & pf_0 \\ pf_1 & (1-p)f_0 \end{vmatrix} \tag{25}$$

If we now implement a coarse graining by integrating over every other time step we arrive at a coarse grained transfer matrix,  $W'$ , that satisfies

$$W' = W^2 \tag{26}$$

The most natural reparametrization is that of the original, *i.e.* we parametrize  $W'$  as

$$\mathbf{W}' = \begin{vmatrix} (1-p')f'_1 & p'f'_0 \\ p'f'_1 & (1-p')f'_0 \end{vmatrix} \tag{27}$$

The equation  $W' = W^2$  then leads to four equations with only three parameters -  $p'$ ,  $f'_1$  and  $f'_0$  - that may be adjusted. One finds that the equations cannot be solved. The interesting conclusion in this case is that the system is ‘‘non-renormalizable’’ with respect to this reparametrization. The intuitive reason why the reparametrization doesn’t work is that the selection makes the ‘‘effective’’ mutation rate from allele one to allele zero different to that from allele zero to allele one. In other words asymmetric hopping probabilities between the two alleles are required. Introducing, such probabilities,  $p'_1$  and  $p'_0$ , leads to the following relations.

$$f'_1 = (1-p_1)f_1^2 + p_1f_0f_1 \tag{28}$$

$$f'_0 = (1-p_0)f_0^2 + p_0f_0f_1 \tag{29}$$

$$p'_1 = p_1 \left( \frac{(1-p_1)f_1 + (1-p_0)f_0}{(1-p_1)f_1 + p_1f_0} \right) \tag{30}$$

$$p'_0 = p_0 \left( \frac{(1-p_0)f_0 + (1-p_1)f_1}{(1-p_0)f_0 + p_0f_1} \right) \tag{31}$$

These relations generally relate the parameters after  $n + 1$  iterations of the RG map, associated with a system of  $t/2^{n+1}$  “degrees of freedom” (time steps), to those after  $n$  iterations corresponding to a system of  $t/2^n$  degrees of freedom. The interpretation of these results is via the “Law of Corresponding States” - where we can see that the dynamics of this one gene genetic system with fitness landscape  $f_1, f_0$  and mutation rates  $p_1$  and  $p_0$  evolving to time  $t/2^n$  is identical to that of a one gene system with fitness landscape  $f'_1, f'_0$  and mutation rates  $p'_1$  and  $p'_0$  evolving to time  $t/2^{n+1}$ . By mapping the system to a smaller and smaller number of degrees of freedom one hopes to realize the RG goal of reaching a system with a sufficiently small number of degrees of freedom that a simple approximation technique may be used. The Eqs. (28-31) compose an exact closed form RG.

We will now consider the asymptotic behaviour of these equations starting with their fixed points. We are interested in the fixed points of (26), i.e.  $W^* = W^{*2}$ . If  $W$  is invertible then this reduces to  $W^* = 1$ . The only solution of this equation is  $p_1 = p_0 = 0, f_1 = f_0 = 1$ . Actually, due to the invariance of the dynamics under any constant rescaling of the fitness values,  $f_1 = f_0 = C$ , where  $C$  is any positive constant, exhibits fixed point behaviour. An appropriate way to take this into account is to consider the quantity  $H = \ln(f_1/f_0)$  as this is automatically invariant under any rescaling. This quantity, in fact, intuitively, plays the role of a magnetic, or ordering, field as it breaks any symmetry between the 1 and 0 alleles. Large  $|H|$  corresponds to a regime of strong selection pressure where one allele is strongly favoured over another. In terms of  $H$  the fixed point is  $H = 0, p_1 = p_0 = 0$ .

If we also consider solutions where  $W$  is not invertible we find the additional solutions:  $H = \infty, p_1 = p_0 = 0$  and  $H = C, p_1 + p_0 = 1$ , where  $C$  is any positive constant, the latter corresponding to a line of fixed points. These fixed points are analogous, as one might expect, of what is found for the one-dimensional Ising model.  $H^* = 0, p_1^* = p_0^* = 0$  is the “ferromagnetic” fixed point corresponding to a critical temperature  $T = 0$ . On the other hand we find a “paramagnetic” line of fixed points at  $p_1^* + p_0^* = 1$  and a “fully aligned” or “frozen” infinite field fixed point at  $H^* = \infty, p_1^* = p_0^* = 0$ .

We can linearize the RG equations in the vicinity of these fixed points. In the vicinity of the ferromagnetic fixed point we have

$$\delta H' = 2\delta H \tag{32}$$

$$\delta p'_1 = 2\delta p_1 \tag{33}$$

$$\delta p'_0 = 2\delta p_0 \tag{34}$$

Thus, we see that the eigenvalues of the RG linearized around this fixed point are all positive, meaning this fixed point is unstable to linear perturbations in both the mutation rates and selection strength.

Considering now the “infinite temperature” paramagnetic fixed points: on the line  $p_1^* + p_0^* = 1$  we have

$$H' = H \tag{35}$$

hence any value of  $H$  is a fixed point, thus confirming that we have a line of fixed points. To illustrate their nature with respect to perturbations in the mutation rate it suffices to consider the case where  $p_0^* = p_1^* = 1/2$  and symmetric perturbations with  $\delta p_0 = \delta p_1 = \delta p$ . In this case, for any  $H$ ,

$$\delta p' = -2(\delta p)^2 \tag{36}$$

so we note that mutation is an irrelevant operator for the paramagnetic fixed points, though, interestingly, the irrelevance only appears at quadratic order. Thus, the RG flows for this system flow from the doubly unstable ferromagnetic fixed point to the line of paramagnetic fixed points. Starting with a system at time  $t$ , with a given selection pressure and mutation rate, the RG maps to a system with stronger selection pressure and higher mutation rate. i.e. the RG takes us away from the critical point.

It should be noted that these fixed points are fixed points for any value of the initial population probability distribution,  $\mathbf{P}(t)$ . In the case of the ferromagnetic fixed point any point within the population simplex is already a fixed point in the infinite population limit. In this sense correlations in the time direction are of long range near this fixed point. For the frozen fixed point the vertex of the simplex of highest fitness is the population attractor while for the paramagnetic fixed points the center of the simplex is the attractor. So the fixed point probability distributions are:  $P_1 = 1, P_0 = 0$ ;  $P_1 = 0, P_0 = 1$ , and  $P_1 = 1/2, P_0 = 1/2$ . Although these are the only fixed point distributions with respect to RG transformations they are not the only fixed points of the dynamics. These may be obtained directly from (17) setting  $\mathbf{P}(t + 1) = \mathbf{P}(t) = \mathbf{P}^*$  for our simple one gene model. One finds for the case  $p_1 = p_0 = p$

$$P_1^* = \left[ ((f_1 - f_0) - p(f_1 + f_0)) \pm ((f_1 - f_0)^2(1 - 2p) + p^2(f_1 + f_0)^2)^{\frac{1}{2}} \right] / 2(f_1 - f_0) \tag{37}$$

which is the “Quasi-species” [5] for this simple model and corresponds to the asymptotic population distribution. Under a coarse graining the Quasi-species distribution is not invariant being mapped into the asymptotic distributions  $P_1 = 1$ , for zero mutation, or  $P_1 = 1/2$  for non-zero mutation.

So what are the “biological” interpretations of these fixed points? The fixed point  $H^* = \infty, p_1^* = p_0^* = 0$  corresponds to the strong selection limit, where the positive selection associated with one allele over another is so strong that the population completely orders after one generation. The ferromagnetic system corresponds to the limit of small selection pressure and small mutation rate so that the population does not “evolve” but remains relatively static. Of course, in the real world, populations are not infinite and “finite size” effects have to be considered. In this case there will be a “spontaneous symmetry breaking” such that the population becomes uniform consisting of either all alleles 1 or alleles 0, the timescale of the symmetry breaking depending on  $1/n^{1/2}$ , where  $n$  is the population size. The paramagnetic fixed points are associated with a strong mutation regime

where selection essentially does not play a role. The strong selection/weak mutation regime one may term “Darwinian” due to its emphasis on strong positive selection, whereas the weak selection/strong mutation regime we may term “Neutral” after the Neutral theory of evolution [14].

With the behaviour in the vicinity of the RG fixed points in hand we may calculate quantities such as the correlation length (time), etc.

## 7. Conclusions

In this paper we have shown how coarse graining is a very natural tool to use when examining the dynamics of genetic systems. In particular, we showed how both meanings of coarse graining: i) writing an effective theory in terms of more appropriate effective degrees of freedom and, ii) explicitly integrating out degrees of freedom naturally made their appearance.

We saw that coarse grainings could be implemented as coordinate transformations or as projections on the configuration space. However, integrating out degrees of freedom is always associated with a projection. Different coarse grainings were seen to be naturally associated with the different genetic operators, a particularly interesting one being the map to the BBB in the case of recombination. We showed that both coarse grainings in space and time naturally showed a

semi-group structure associated with the RG.

We also presented for the first time an explicit exact RG for a concrete simple genetic system - a one gene, two allele model evolving under the action of mutation and selection - showing how an explicit coarse graining RG led to identical dynamics for different “corresponding states” explicitly showing the relationship between the renormalized parameters of these different “corresponding” systems. We examined the fixed points of the RG finding two isolated fixed points associated with the strong selection/zero mutation and zero selection/zero regimes respectively, and a line of fixed points associated with the weak selection/strong mutation regime. By considering the RG in the neighbourhood of these fixed points we showed that the weak selection/strong mutation regime was the only stable regime in terms of perturbations in the selection pressure and mutation. We strongly believe that the RG has an important role to play in developing a more quantitative understanding of the dynamics of genetic systems.

## Acknowledgements

The work of CRS was supported by Conacyt grants 30422-E and 41221-F. CRS is grateful to Riccardo Poli, Alden Wright and Michael Vose for useful conversations. AZ is grateful for the hospitality of the Instituto de Ciencias Nucleares, UNAM.

- 
1. C.R. Stephens, *Acta. Phys. Slov.* **52** (2002) 515.
  2. B. Drossel, *Advances in Physics*, to be published; preprint cond-mat (2001) /0101409.
  3. R. Bürger, *The Mathematical Theory of Selection, Recombination and Mutation*, Wiley Series in Mathematical and Computational Biology (2000).
  4. D.R. Nelson and M.E. Fisher, *Annals of Physics* **91** (1975) 226.
  5. M. Eigen, *Die Naturwissenschaften* **10** (1971) 465.
  6. C.R. Stephens and J. Mora, *Genetic Programming and Evolvable Hardware* **2** (2001) 7.
  7. P.F. Stadler and C.R. Stephens, *Comments on Theoretical Biology* **8** (2003) 1.
  8. C.R. Stephens and H. Waelbroeck, *Effective Degrees of Freedom in Genetic Algorithms and the Block Hypothesis, Proceedings of the ICGA7*, (ed. T. Bäck, Morgan Kaufmann, San Mateo 1997) p. 34.
  9. C.R. Stephens and H. Waelbroeck, *Physical Review* **D57** (1998) 3251.
  10. C.R. Stephens, “Some Exact Results from a Coarse Grained Formulation of Genetic Dynamics”. In L. Spector *et al* (*Proceedings of GECCO 2001* Morgan Kaufmann, San Francisco, 2001) p. 631.
  11. M.D. Vose, *The Simple Genetic Algorithm: Foundations and Theory*, (MIT Press, Cambridge MA, 1999).
  12. A. Aguilar, J. Rowe, and C.R. Stephens, “Coarse Graining in Genetic Algorithms: Some Issues and Examples”, In. E. Cantú Paz *et al* eds. *Proceedings of GECCO 2001*, (Springer-Verlag, Berlin, 2003).
  13. I. Leuthäusser, *J. Phys. Chem.* **84**, (1986) 1884.
  14. M. Kimura, *The Neutral Theory of Molecular Evolution*, (Cambridge University Press, 1983).