Are There Ergodic Limits to Evolution? Ergodic Exploration of Genome Space and Convergence

Tom C. B. McLeish

Departments of Physics and Chemistry and Biophysical Sciences Institute,
Durham University

Abstract

We examine the analogy between evolutionary dynamics and statistical mechanics to include the fundamental question of *ergodicity* – the representative exploration of the space of possible states (in the case of evolution this is genome space). Several properties of evolutionary dynamics are identified that allow a generalisation of the ergodic dynamics, familiar in dynamical systems theory, to evolution. Two classes of evolved biological structure then arise, differentiated by the qualitative duration of their evolutionary timescales. The first class has an ergodicity timescale (the time required for representative genome exploration) longer than available evolutionary time, and has incompletely explored the genotypic and phenotypic space of its possibilities. This case generates no expectation of convergence to an optimal phenotype or possibility of its prediction. The second, more interesting, class exhibits an evolutionary form of ergodicity – essentially all of the structural space within the constraints of slower evolutionary variables have been sampled; the ergodicity timescale for the system evolution is less than the evolutionary time. In this case some convergence toward similar optima may be expected for equivalent systems in different species where both possess ergodic evolutionary dynamics. When the fitness maximum is set by physical, rather than co-evolved, constraints, it is additionally possible to make predictions of some properties of the evolved structures and systems. We propose four structures that emerge from evolution within genotypes whose fitness is induced from their phenotypes. Together these result in an exponential speeding up of evolution, when compared to complete exploration of genomic space. We illustrate a possible case of application and a prediction of convergence together with attaining a physical fitness optimum in the case of invertebrate compound eye resolution.

1. Introduction

For several generations of thinkers in the field of evolutionary dynamics, there has been a fruitful conversation with the concepts and methodologies of statistical mechanics[1]. The analogy arises because random mutation between alleles at the genotype level induces a coarse-grained diffusion within the space of coded structures at the phenotype level, in a similar way that inter-microstate dynamics generates the sampling of macrostates in statistical mechanics. So divergence among genotypes (*e.g.* in bacteria) may nonetheless map onto a convergence in phenotype, in a manner isomorphic to the mapping of large numbers of configurational microstates into the same macrostate in statistical mechanics. Fisher[2] and Wright[3] were early to spot some classes of derived mathematical model that present themselves as analytical tools for such stochastic dynamics of populations. There are three principal common ingredients that make the analogy between statistical mechanics and evolution fruitful: (i) a very large space of states; (ii) a coarse-grained set of properties that emerge from the microscopic states; (iii) a stochastic dynamical process that moves the system from one state, or set of states, to another. In statistical mechanics, the large underlying space is that of the system phase space (the 'giant' space formed by all of the degrees of freedom of all of its components), the emergent coarse-grained structure is the physics of 'macrostates' (such as defined by pressure or volume), and the stochastic dynamics are generated by random thermal

motion of all the degrees of freedom. In evolution, those structures map onto (i) the genotype, (ii) the phenotype and (iii) the processes of random mutation respectively[4].

Further analogies, together with their mathematical isophorphisms, present themselves. The macroscopic consequence of random dynamics among microstates in statistical mechanics is the optimisation of an appropriate functional – termed the 'free energy', F, whose exact form depends on the external constraints on the system. A macrostate is sampled with probability $\varphi(F)$ proportional to the 'Boltzmann weight' $\varphi(F) \approx \exp(-F/k_B T)$, where T is the absolute temperature. The analogy in evolutionary dynamics is with a quantity that can be called 'free fitness' Φ. As proposed by Yoh Iwasa[5], demonstrated by Sella and Hirsh[6] in the general case, and Berg and Lässig[7], for the case of transcription factor binding, there is an exact mathematical correspondence between the Wright–Fisher process and the canonical ensemble of statistical mechanics in steady state[8]. The ratio of forward and reverse fixation rates $\varphi(\delta F)/\varphi(-\delta F) = \exp(\nu \delta F)$, for example, implies that the probability distribution of the most recent common ancestor is a Boltzmann distribution with an effective temperature $T_{eff}=(2(N-1))^{-1}$ for haploid asexual populations. As shown by Sella and Hirsh, this means that in steady-state evolutionary processes, there is a balance between mutational entropy S and mean fitness $<F>$, where a quantity analogous to the Helmholtz free energy, the free fitness

$$\Phi = <F> + T_{eff}\, S \qquad\qquad (1)$$

is at maximum. In just the same way that the free energy 'folds in' the entropy of its macrostate – *via* the collection of microstates that contribute to it, so the free fitness folds in the genotypic entropy – the ensemble of genotypes that correspond to a single phenotype. So the emergence of populations of organisms with many different genotypes but exhibiting a common phenotype corresponds to a statistical mechanical system with a high entropy (increasing the likelihood of its realisation in dynamic equilibrium). A phenotype may be selected, not because it has the highest fitness, but because it corresponds to a large number of genotypes, while possessing an acceptable fitness.

Although the Wright-Fisher process applies to asexual reproduction, the concepts of genotypic entropy, induced fitness on the genotype, and in particular the dynamics of genome exploration, sustain analogies under sexual reproduction[1] as well.

In this paper, we take the analogy between statistical mechanics and evolution one step further and examine the correspondence between one of the underlying assumptions. Without these foundational assumptions we are not able to define and write down quantities such as (1) in the first place, in either of the two cases. Of these, the most fundamental is the *ergodic assumption*, originally due to Boltzmann[9]. This simply states that the time average of a system in thermodynamic equilibrium is equivalent to a phase space average. However, it goes beyond the simple statement of the relative probabilities of occupation of different microstates at equilibrium to a statement about the nature of the *dynamics of the trajectory* by which these microstates are visited. Boltzmann realised that a set of properties must be true of the trajectory in order for the assumptions of equilibrium statistical mechanics to be valid. To see this, we consider an equivalent statement of the ergodic hypothesis: it assumes that the dynamics are sufficiently random such that, after long enough times, the time spent in any region of state space is simply proportional to its volume (strictly we should say 'hypervolume' as the space of possible states has very many degrees of freedom, so is a very high dimensional space). Note that this assumption, while simple to state, is by no means obviously or trivially realised – the number of microstates in a macroscopic system where the mean energy content per degree of freedom is constant, grows super-exponentially with the number of degrees of freedom, N, as N^N. This means that, for all practical purposes, the fraction of all microstates actually explored in any macroscopic observation is exponentially small. The region of state space sampled in ergodic dynamics must however pass close to every point in the full state

space, while occupying negligible (hyper)volume.   This is possible if the local dimensionality of the trajectory is much smaller than the dimensionality of the full space. In statistical mechanics, even though only a vanishingly small fraction of possible states of a system are actually sampled in any measurement, yet the measured macroscopic properties are such that it would display if this were true. This includes the attainment of thermodynamic equilibrium (we note that this differs from the simpler concept of static mechanical equilibrium – in thermodynamic equilibrium the occupation of microstates is continually in flux), and the corresponding minimisation of free energy.

Our purpose here will be to lay out the broad questions and consequences of an analogy to ergodic dynamics in random evolution. Is there a useful analogy to ergodicity for the dynamic exploration of the space of all possible genomes? Just as in the statistical mechanics of any macroscopic system, the number of all possible 'states' (genomes) is so enormously large that there is no hope that any evolving population of organisms, no matter how simple or numerous, could have explored it in entirety in its population history. Yet, if the structure of a dynamical system has the correct properties, it is possible that the state space *is* explored in a *representative* way within an attainable ('ergodic') timescale τ_e. If this can be shown to be true for the exploration of genotype state space within evolutionary dynamics, then the consequences are very significant. In particular, if the evolutionary ergodic search time of a genome subspace for any corresponding phenotype can be calculated, then at evolutionary timescales longer than this, we expect that a fitness optimum can be found, if one exists. This would provide a conceptual basis for understanding convergence in evolution, and for differentiating between systems that converge (onto optima) and those that do not.

The paper is laid out as follows. In section 2, we develop the ingredients that a definition of ergodic evolution would require, and what properties an ergodic evolutionary dynamics might display. In section 3 we then explore in detail four specific properties of genotype-phenotype mapping that contribute to an ergodic exploration of genomes. In section 4 we lay out a research programme into one specific example candidate phenotype for ergodic evolution of its genotype (the acuity of insect compound eyes), referring to data on current and archaic species, and on what is known about the relevant gene control network.

2. Ergodicity in Evolutionary Dynamics

In a high-dimensional dynamical system, ergodicity is achieved by the structure of the path of the system state-point through its state space. Broadly, the system never returns to a point previously visited (if it did that its dynamics would simply be periodic in the deterministic case). Neither can it be guaranteed to visit every state (there are far too many). However, for any small distance required, there is a dynamical time beyond which it will have visited as close, or closer than, that distance to any state point. We will explore here whether the space of possible (and viable) genomes also possesses this property under evolutionary dynamics.

A three-dimensional illustration, without the complication of dynamics, is furnished by a polymer gel. The polymer chains may occupy only a tiny fraction of the volume of the gel, yet dominate its properties over the solvent (creating a solid rather than a liquid). They accomplish this through their locally 1-dimensional form, embedded within the 3-dimensional solvent. By analogy, the trajectory of a thermodynamic system within phase space necessary to the establishment of thermodynamic equilibrium needs to possess a 'threadlike' geometry. By this we mean that it is (i) connected; (ii) locally of lower dimensionality than the full phase space, (iii) embedded within the full phase space. By this means it is possible for each important point in phase space is 'close' to a real, sampled trajectory, while the trajectory itself fills a vanishing fraction of the phase volume. Unsurprisingly, the cases of dynamical systems for which ergodic behaviour has been proved are both few in number and contrived in form. One example is the simplest space with negative

Riemannian curvature, first investigated by Artin[10], another the 'Lorenz gas' – a square two-dimensional domain with a reflecting circle removed from its centre, shown to be ergodic by Sinai[11]. There are general considerations that suggest the application of the ergodic assumption for large material systems. However, it remains the case that the foundation of statistical mechanics for the general case is largely a matter of informed and motivated assumption.

As a strong warning against a blanket application of the ergodic principle, there are many examples of physical systems where equilibrium statistical mechanics does not hold, essentially because of the breakdown of ergodicity. These are known as 'glassy'. The energy landscape of a glass is such that the thermally activated dynamics of its degrees of freedom are unable to drive it to all regions of its phase space within the timescale of observation. This has real, physical consequences, such as the emergence of a solid elastic modulus in systems that possess the same symmetries as equilibrium fluids.

Furthermore, all systems, when sampled on small enough timescales, fail to demonstrate equilibrium properties in this way; the question is simply the ratio of the observation timescale τ_o to an *ergodic timescale* τ_e, which identifies the minimum time necessary for the uniform exploration of phase space. So a second case in which ergodicity does not hold is during the dynamical approach towards equilibrium. The ergodic timescale τ_e can equivalently be thought of as the time taken by the system to reach equilibrium. The 'ergodic ratio'

$$R_e \equiv \tau_o / \tau_e \qquad (2)$$

defines the applicability of statistical mechanics, and its consequences of a dynamic convergence from any starting condition to equilibrium. Equilibrium statistical mechanics makes predictions for even very complex systems when they are at equilibrium, for in this case the observed macrostate will be the one that minimises the free energy F. But all of this powerful predictive machinery is lost if $R_e < 1$.

The same questions of timescale and structure of phase space arises in evolution. Just as in the case of statistical mechanics, the number of genotypes in principle possible in even the simplest organisms is super-exponentially large[12]. For example, the size Ω of genotype space for a protein consisting of N residues chosen from an alphabet of amino acids of size m is

$$\Omega = m^N \qquad (3)$$

Dryden *et al.*[13] point out that even in the case of the high evolutionary rates of bacteria, their large population sizes and their available evolutionary timescales of 100 Myears, there is a strong limit on the size of genome of a single protein that can be explored in its entirety, even when the effective value of m is reduced by homology. For example, in the case of a 100-residue bacterial protein, the population size of possible sequences with 20 amino acids to choose at each residue, is 20^{100} (approx. 10^{130}). Using upper limits to bacterial genome mutation rates, the global numbers of bacteria, and a longest timescale available to evolution of 4 Gyr gives an upper estimate of the number of potentially explored sequences as 10^{43}, a tiny fraction (10^{-87}) of the possible sequence space. The situation is of course much worse for longer proteins, or for systems of several proteins. Just as in statistical mechanics, it is the high dimensionality (in this case, the sequence length) that is responsible for the enormous size of the full sequence space.

But this does not mean that S J Gould's hypothesis[14], that the evolved trajectory of life is necessarily contingent, follows, any more than would the corresponding notion in statistical mechanics, that it is impossible to entertain any notion of predictable equilibrium. If I perform an experiment of releasing a mole of nitrogen molecules at 300K into an evacuated chamber 1m^3 in volume, within a few hundredths of a second the gas will assume an equilibrium pressure close to the value 300R (Where R is the ideal gas constant). This will happen no matter how many times I

repeat the experiment even though the trajectories of individual molecules are completely different on each occasion. This is true because of the ergodic property of those trajectories – they cannot visit all of state space, but they can explore it representatively. In so doing, they bring the system *convergently* (across the many realisations) into thermodynamic equilibrium, where the free energy is minimised with respect to the constraints (in this case the imposed volume), and the predictive ideal gas law, *pV=nRT*, observed.

In a similar way within evolutionary systems, for 'convergent evolution'[15] to be realised, this requires the corresponding dual properties of an attracting optimal macrostatestate (the observed phenotype) and the predictability of that state (the one that maximises fitness under the effective constraints). This can be realised if exploration of genomic space is representative rather than complete, in just the same way that the gas molecule trajectories are representative rather than complete. This is what is meant by *ergodic* evolutionary dynamics. To make this insight a little more quantitative, we need to show that the exponentially large hypervolume represented by equation (3) above can be reduced to a much smaller space (ideally growing in an algebraic, rather than exponential way, with system size). We also need to show that this space explores and connects the larger space in a representative way. Such an achievement wins a double reward: first we will be able to classify evolutionary genotype systems and their phenotypes into those whose evolutionary space is ergodic (those for which the evolutionary time – now the interpretation of τ_0 in equation 2 – is greater than an ergodic time for genotype exploration, τ_e), and those for which it is not; second, for those in which ergodicity holds we may be able to make real predictions[16] of the optimal phenotype in some cases.

We need to note other assumptions that need to be made in order to apply equation (2) to the case of evolution. One is that the rate of adoption or rejection of mutation is much slower than the mutation rate itself. This allows the fitness landscape to impact on the dynamics of its exploration. It will be necessary, for example, for the effect of deleterious mutations to 'channel' the pathways of genotype trajectories (discussed in 3.1 below). In the example of the asexual Wright-Fisher process this is guaranteed by a single 'step' in genome space being the adoption of a mutation by the entire population. But in other processes of incomplete adoption there is nevertheless an effective population number that assures this separation of timescales. A second separation of timescales is that of the evolving (and possibly ergodic) subsystem and the environment around it that sets its fitness landscape. When the environment is set by other evolved species, this is not necessarily true and needs to be checked on a case-by-case basis.

Even when these assumptions are satisfied, and an ergodic evolutionary trajectory can be defined, the idea that one might then be able to predict convergence onto an optimal phenotype appears at first sight to be a very bold claim. In many cases we might expect the boundary conditions for a optimally-evolved subsystem themselves to be contingently set by a larger, non-ergodic system. For example, the neck length of a herbivorous mammal may be optimally evolved, but contingently so depending on the height of the vegetation it feeds on, and on its own inherited body plan. These constraints will have, in turn, evolved contingently within larger, and different parts of the universe of genomes. However, some phenotypes possess fitness landscapes that depend on physical, rather than biologically evolved, constraints. In these cases the identification of ergodic evolutionary dynamics can in principle lead to predictability. In cases for which several different species possess the analogous subsystems (eyes, for example), which are themselves subject to physically-defined optima, and have evolved ergodically, we may also be able to explain convergence of their phenotypes in this way. We will consider an example of this type in the final section.

The first task is to show how the exponentially large genotype space may be explored representatively, and in what sense the ergodic hypothesis can be applied to understand evolutionary dynamics.

3. Ergodic Structures of Genotype-Phenotype Dynamics

Reasons for a large reduction in the actual reduction of explored genotype space, without restriction on actualised phenotypes, have been pointed out by Gavrilets and others[17]. However, we here attempt to bring them together into a unified attempt at outlining how ergodicity might be redefined in the case of genotype exploration, and what structures of evolutionary dynamics might replace those on which dynamical systems theory relies for ergodicity. Energy and momentum conservation, deterministic nonlinear equations of motion and proofs of divergence of local trajectories, normal within 'ergodic theory' are all unavailable to us. Fortunately other properties of evolutionary dynamics enter to take their place.

Fortunately there are, in their place, a number of other properties of evolutionary dynamics, which present as candidates to take their place. First, the 'threadlike' nature of realistic exploration of genomic space is guaranteed by: (i) the tiny fraction of mutations from any viable mutation that is itself viable (even with reduced fitness), and (ii) the way mutation occurs in *populations*, themselves carrying a range of alleles for each gene, so permitting the simultaneous exploration of mutational space in the immediate vicinity of the current genotypes. The genotypic space may, as a result, contain 'holes' at all dimensions which can never be explored since their phenotypes are not viable. Second, the restriction of evolutionary dynamics to the pathways of viable mutations creates a network that is able to span exponentially larger hypervolumes of genome space than it connects within its own volume. Third, the special properties of evolved genomes are likely to have created nested subspaces of increased complexity that may be searched hierarchically. Fourth, the actual generalised 'distance' between genomes (measured in terms of the number of loci requiring mutations to reach one from the other) remains algebraic in the system size in spite of the exponential size of its hypervolume.

We next examine each of these accelerating features of the search in genotype space in a little more detail.

3.1 Pathway Confinement
The confinement of (otherwise) random walks by a fitness/free-energy landscape arises in polymer physics[18]. We can ask about the typical fluctuations in the pathway of a random walk r(s) of step length b, (a polymer) in a landscape in which it is coupled to random fluctuations in a field $\phi(r)$ that predisposes the polymer to follow valleys in the field. The connection with a pathway in evolution is that the polymer becomes a pathway through genome space, and the field $\phi(r)$ describes local fitness variations that fluctuate across the genome space (on the scale of a single mutation). The property of confinement of paths subject to a random potential is a surprising one: although the random field has a zero mean, it always serves to confine, or contract, the polymer (or path) laterally to a narrow 'corridor', whose width depends on the structure of the random potential. The relevance of this effect is clear in the case of an evolutionary path, if the polymer analogy holds: unfavourable regions close to a path between local minima keep the path form exploring large regions of the space which it would otherwise diffuse into.

To make this notion more quantitative, we can specify the statistics of the random field by a localised Gaussian variance of w (this sets its strength) and localisation length κ^{-1} (this sets its range) within a suitable modal function for localisation) such that:

$$\langle \varphi(r')\varphi(r'') \rangle = w \frac{e^{-\kappa|r'-r''|}}{|r'-r''|} \tag{4}$$

We can couple this noise in the energy of a polymer landscape (or equivalently the fitness landscape of an evolutionary walk) *via* the total Hamiltonian for both noise and polymer fields, in Fourier transform (with variable k the transform of r):

$$H = \frac{1}{2}\sum_k \{k^2 \tilde{\rho}_k \tilde{\rho}_{-k} + \tilde{\varphi}_k \tilde{\rho}_{-k} + \tilde{\rho}_k \tilde{\varphi}_{-k} + w^{-1}(k^2 + \kappa^2)\tilde{\varphi}_k \tilde{\varphi}_{-k}\} \tag{5}$$

Here $\tilde{\rho}_k$ is the Fourier transform of the polymer (walk) density and $\tilde{\varphi}_k$ that of the fitness field. In this form it is simple to integrate out the possible configurations of the random field (taking a Gaussian weight for its distribution). This then leaves an effective correlation expression for the polymer/path alone. If in addition we impose constraints on the two ends of the walk (in the polymer case this is fixing the two ends of the chain, in the evolutionary pathway case it is asking about the likely configuration of the mutation pathway between two genomes) then we find that the typical excursion of the path perpendicular to the lowest energy (highest fitness) pathway is confined to a corridor whose width inversely of the height w, of the landscape fluctuations as $a = \kappa w^{-1/2}$. At the limit of high fluctuations this cut-off by a single mutation/step width in the high-w limit (see figure 1).
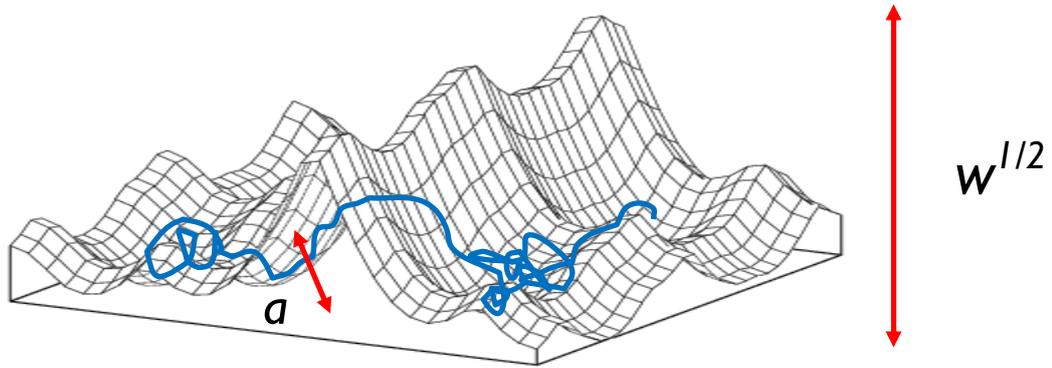

Figure 1. An evolutionary pathway as a Gaussian polymer in a rugged landscape: width confinement around the highest fitness trajectory scales as the inverse of the fitness ('energy') fluctuations.

To gain a hold on suitable quantitative values to choose for the range (in Hamming distance) and strength (in fitness) of the model random potential, we need to map this simple model onto data for fitness variations for a particular example. One potential source of data is provided by studies on DNA sequence polymorphism. One example, comparing two Drosophila species to extract the distribution of deleterious mutations[19] concluded that the probability of a neutral mutation at a single codon was less than 20%, and that the mean fitness change s was such that $sN_e>1$ (N_e is the effective population size). This signifies that the effective 'fitness noise function' $\phi(r)$ in our simple model should, in this case, be chosen so that the range is that of a single mutation (although the finding of reference 19 that the distribution of fitness variation was log-normal rather than normal provides a caveat to the very approximate nature of this model).

The conclusion from the best choice of model parameters given mutation/fitness data is that evolutionary paths on landscapes with fluctuating fitness are in general confined to the 'most neutral' trajectory very closely. If the path length (or Hamming distance) between the start and end points of the walk is m', and the dimension of the space N, (following our earlier definitions relating to amino acid alphabet and genome length), then the explored volume of genome space after confinement is reduced from m'^N to $m'^2 a^{(N-1)}$. Note that this is not a special property of genome

spaces as such, but emerges from the stochasticity of *both* the path *and* the potential of fitness fluctuations in which it is located.

3.2 Percolation Structure

We next need to ask about the connectedness of the threadlike paths created by fitness landscapes. How often do they branch? What fraction of the total genome space to they occupy? An important question is about their effective dimension – how many sites on the connected paths lie within *n* steps of a given site? The equivalent phenomenon in the statistical mechanics of networks is *percolation*[20]. An abstract lattice of nodes connected by links possesses a remarkable property of connectivity when links are given a stochastic property of 'occupation' with a given probability *p*. If the functionality (the number of nodes connected to a given node) of the network, *z*, is large then there is a (small) critical value of *p*, $p_c \cong (1/z)$ above which there exists at least one connected cluster which spans the entire space. The distribution of sizes of the individual clusters (apart from the effectively infinite one) is given by a power law. But more importantly the structure of the individual 'percolation cluster' is itself *fractal* – the number of occupied links ω accessible within a connected pathway of length *s* from a given link grows as a power-law, not exponentially, with *s*. When *z* is large the power is close to 2 so that

$$\omega \cong s^{d_s} \quad \text{with } d_s = 2 \tag{6}$$

d_s is termed the 'spectral dimension' of the fractal. The random walk property of the connected paths within clusters themselves (they are Gaussian random walks with fractal dimension 2) implies that the spatial fractal dimension of the percolation clusters (in simple terms how their mass scales with their size) in this near mean-field limit is effectively 4. So in terms of the span *r* of the clusters within the (Euclidian) distance measure of the network,

$$\omega \cong r^{d_f} \quad \text{with } d_f = 4 \tag{7}$$

The consequence of the power-law volume of these connected clusters of links within a hyperspace whose volume increases exponentially with size, is that they contain unoccupied 'holes' of all size scales.

That it is possible to map these percolation clusters onto evolutionary dynamics is the hypothesis of Gravner *et al.* [21]. The idea here is an extension of the polymer physics analogy – that an evolutionary pathway can be thought of as a connected polymer embedded in a much larger space – introduced in section 3.2 above. The motivation of the narrowness, the polymer-like character, of the pathway follows, as before, from the observation that most steps in genome space (most mutations to the genome) are deleterious and will not be adopted by the population. However, the extension to percolation clusters is effectively saying that the correct polymer analogy is with *branched polymers*. Although most mutations are not sustained along the evolutionary track of a genome, there is more than one 'direction' in which the pathway can find viable genomes from at least some points. In polymer physics, there are typically separate, disconnected branched polymers made this way. These would correspond to regions of the genome space that can be mutually accessed by point mutation. The other property of polymeric percolation cluster is is we have seen that, above a critical probability for a node to be occupied, there is at least one cluster that connects across the entire space, while only occupying a tiny fraction of its volume. This idea has clear implications, if relevant, for ergodic exploration of genome space.

The degree of branching of a polymer in a percolation cluster is controlled by two parameters: (1) the number of possible connections to each node, or 'functionality' *z*; (2) the probability that

polymer occupies these branches. Gravner *et al.* additionally observe that the effective functionality *z* of a high-dimensional genome network is greater than the relatively low number of new genomes linked to a single genome by a single mutation (this is *m-1*). The effective value of *z* is a much larger number. This is because the nodes of the network are populations of genomes that differ in the *n* qualities of a phenotype. This delivers the much larger effective functionality 2^n.

However, the percolation hypothesis comes at a price. If the dimension of the genome space really sustains, throughout its domain, the low-dimension (4) of an ideal percolation cluster, then there are three difficult issues that remain to be addressed. The first is that we now have to explain why the system should be driven to the special state of criticality in which the link occupation probability (which is a measure of the fitness likelihood of the phenotype of a random genome) should be linked to the (inverse) number of neighbouring genomes accessible by a mutation. The second is, as Gravner *et al.* point out, that the occupation probability of a bond is not a uniform number, but rather subject to the large 'fitness noise'. We saw above in 3.1 that this tends to confine trajectories to as neutral a path as possible. The third problem is that as the dimension of the embedding space increases, so the critical percolation cluster is able to explore a smaller and smaller volume. There is only one spanning cluster, together with many smaller clusters (in fact a whole power-law distribution of them) which do not span the whole space, and are disconnected from the single spanning cluster. Although the necessary dimensional reduction has been achieved (from the vast space size of equation 3 to a subspace of dimension 4), too much of the space is now inaccessible. Fortunately there is another more subtle way of exploring high-dimensional spaces that retains the ability to search efficiently together with an accessibility of all regions in the space.

3.3 Nested hierarchical subspaces

Evolution is not the only problem in biology that may involve search for optimal points in very high dimensions. Another is the search for target operons and other binding sites by DNA-binding proteins. This was famously analysed first by Adam Delbrück as an example of a requirement of specific structure on a search space (in this case that of dimensional reduction, closely parallel to the nested structures of genome spaces we discuss below)[22]. The celebrated problem of protein folding[23] presents the same challenge of a search within an exponentially large state space. In folding a protein comprised of *q* amino acid residues to find its globular native fold, it needs to identify one among the 3^q possible rotational isomeric states of the polypeptide chain (this rough estimate counts overlapping configurations, but the non-overlapping constraint still gives an exponentially large number of configurations to search through). Furthermore, a protein needs to locate the single configuration corresponding to its native fold within biophysical timescales.

A key 'benchmark' result for all processes that require a search in high dimensional spaces, which is easy to formulate, calculate and state, is the mean first-passage reaction-rate for a diffuser, released at the boundary of a hypersphere of radius *R*, in *d* dimensions, to arrive at a target of radius R_N at its centre. This is a fundamental result for all random searches in high dimensional spaces, including both protein folding and genotype exploration. In the case of protein folding, the 'dimensions' or 'degrees of freedom' are the rotational states at each C_α carbon. The diffusive process models the random thermal switching between these states. In the case of evolution each dimension represents the choice of base pairs for one codon. The diffusion in the high dimensional space represents in this case the point mutations of codons. The diffusion calculation makes a continuous approximation for cases of discrete variables in these two cases, for each protein rotational state can take one of three values, and each codon can code for one of 20 amino acids. However, the results do not depend heavily on this continuum approximation. The result (for *d>3*) for the mean search time $\tau_{d\geq3}$, in terms of the diffusion constant *D*, the radius of the target sphere R_N and the ratio of the hypersphere radius to the target radius, $\sigma=R/R_N$ is[24]

$$\tau_{d \geq 3} = \frac{R_N^2}{d(d-2)D} \sigma^2 \left(\sigma^{d-2} - \sigma^{-2} - \frac{d}{2} \left(1 - \sigma^{-2} \right) \right) \qquad (8)$$

This is a very instructive exact result dominated by the first term in the bracket. It says that, in high dimensional spaces, the classic diffusion rate to explore a region of size R in 1 or 2 dimensions, D/R^2 is exponentially slowed down by σ^d. In the context of the protein folding problem, this is the source of the 'Levinthal Paradox' – that no protein could fold if it had to search at random through all its possible conformations. The culprit is not the large size of the search space, nor the small size of the target space, nor the speed of local configuration changes measured by the effective diffusion constant, but purely the high dimensionality, or number of degrees of freedom, of the search space.

In protein folding there is only one solution to this impossible task of searching simultaneously though hundreds of degrees of freedom for the native state – which is not to. Instead, the energetic local interactions of the protein residues, both polar and hydrophobic, constrain the search at any one stage to a small number of degrees of freedom. This is the way to understand the 'folding funnel'. An impossible simultaneous search in a high number of degrees of freedom is replaced by a sequential search in a series of lower-dimensional spaces. The total sum of the dimensions of all subspaces must add to the dimension of the full search space, but dividing the space into such a hierarchy of searches reduces the search time from unimaginable quantities to biophysical timescales. In terms of the notation of (5), a search timescale $\tau \approx \sigma^{(d)}$ is replaced by a sum of sequential searches

$$\tau \approx \sum_i \sigma^{(d_i)} \qquad (9)$$

This accelerated search time is typically dominated by the search in the intermediate subspace with the largest dimension d_i (this is sometimes call the 'transition state' in the protein folding literature). These searches can be marshalled either by non-native interactions providing additional coding to folding pathways, or by chaperone-assisted folding, or both of these[25].

In the case of protein folding, the existence and structure of these intermediate folding subspaces are themselves of course the result of evolutionary pressure on the primary sequence, just as much as is the stability and function of the native fold. In some cases, the existence of secondary structures (helices or sheets) in proteins generate intermediate subspaces in a natural way. For example helices may pre-form, then search in the low-dimensional space of their relative configurations, stabilised by non-native hydrophobic surfaces, for the native fold[18]. There is some evidence, however, that the choice of native interactions that stabilised a single native fold from a large alphabet of residues has the property of inducing a folding cascade around it, as well as the possibility to encode folding strategies within the primary sequence in addition to native stability[25].

The analogy of equation (9) for the mapping onto genome evolution is:

$$\tau \approx \sum_i m^{(n_i)} \qquad ; \text{with } \sum_i n_i = N \qquad (10)$$

where n_i is the number of loci in a genome that may code for a protein residues that together define an aspect of the phenotype. The search time given by (10) is enormously faster than the naïve 'complete-search' time of $\tau \approx m^N$. The question of its applicability to evolutionary spaces themselves naturally arises – does ergodic exploration of genotype space exploit the huge efficiency of hierarchical search structures? At one level the answer is most definitely affirmative – this is an unusual way of thinking about proteins. The very fact that a long genome codes for the synthesis of many individual gene networks and molecules is, from this point of view, a mechanism for reducing

the search time within genome space by exactly this strategy. A particular phenotype is coded for by a small subsection of a genome. If evolution of this section may proceed independently from the rest of the genome, then the search for optima within the space of possible phenotypes may be greatly accelerated. Very rapid and efficient searches may be generated by the independent or subsequent evolution of hierarchical structures within proteins themselves. An example of this strategy will be given in section 4 below.

3.4 Restriction of genome distance

One significant property of very high dimensional spaces is the relative proximity (in Hamming or Euclidian distance) of any two points in the space in spite of the exponentially large volume that they occupy. Along with the mechanisms of dynamical restriction in 3.1 – 3.3 above, this assists the apparently impossible task of ergodic exploration of genome space. For the hypercube of dimension *N* (genome length) and lateral size *m* (amino acid alphabet) with the exponentially large hypervolume of equation (3), the maximal (Euclidian) distance between any two genomes still only weakly and algebraically grows with *N* as $mN^{1/2}$. Alternatively, any genome can be reached from any other by mutating at most *N* residues (the Hamming distance). So, providing that the exponentially large number of stochastic choices that mutations generate can be limited (by strategies 1-3), property-4 assures that a genome representing an optimal fitness can in principle be reached within timescales even as short as the single mutation time (in the limit of directed evolution – an unrealistically extreme, but informative case).

As well as representative exploration of genome space, there are other genetic phenomena that rely on the spatial closeness of genomes, such as the interaction of different mutations (epistatis)[26] and the effect of mutations on more than one phenotype (pleiotropy)[27].

It is possible to estimate the ergodic time of a genomic sub-space exploration from a combination of all four search-reduction strategies. If the largest sub-space of the search is a section of genome length *n*, then the search volume without local fitness fluctuations is m^n. However, as we have seen, the local fitness landscape will localise the space to a threadlike region of width *a*, and length *L*. The segment corresponds to the total length of a segment of percolation cluster, which itself scales as the square of the direct genome distance, so $L \equiv n^2$. The final, reduced genome space that must be sampled for ergodic exploration of the genotype is therefore, up to constants,

$$\Omega_g = a^{(n-1)}n^2 \qquad\qquad (11)$$

If the mutation rate per generation per nucleotide is μ for the species in question, and the population size *N*, then the estimate for the ergodic time (in generations) is:

$$\tau_e \approx (n\mu N)^{-1} a^{(n-1)} n^2 \qquad\qquad (12)$$

This timescale can still be very long compared to a fixation time, depending primarily on the localisation distance *a*, but not unreasonably so, as we will see in the following example. The localisation distance is set by the prevalence of unviable mutations – as most mutations are deleterious, this is short (of order 2), as it needs to be for the timescales of (12) to be controlled.

We may illustrate the power of ergodic genome exploration described by the estimate of (12) applied to an insect species such as *Drosophila Melanogaster* (with an eye on the example in the next section). Global population size is hard to estimate, but a conservative estimate is 10^{10} individuals, with a measured mutation rate per base pair of $\mu=10^{-10}$ per generation (of 30 days)[28]. Setting τ_e to 100Myr permits complete exploration of the viable mutations, and their phenotypes, within the span of a 15-codon subspace. Of course, all subsequences of this length may be explored in the same span of time, as must be the case for the evolution of entire new species on this

timescale, but beyond this length there is no assurance of ergodic exploration. A trait whose value that can be coded for within this length, however, is a candidate for optimisation *via* ergodic exploration within evolutionary time.

4. An Example – Optimising the Invertebrate Compound Eye

One challenge to experimental evaluation of evolution is the very identification of optima. How would we tell when an organism, or a subsystem within that organism, has achieved a maximal fitness? In the majority of cases fitness, though well-defined in terms of production of viable offspring, is measurable only for the current phenotypes, and not for an entire 'landscape' of possible phenotypes. So it is in general not possible to decide whether a current fittest phenotype is locally maximised for fitness, or how close it is to a potential maximum. Within the contingencies of environment, predators and food chains there is in general no adequate model for identifying optima in the genotype-phenotype landscape. We have already shown that, since genotype space is so large, even for systems encoded by a few genes, it is only conceivable that optima in fitness are realised when the associated genome space is explored ergodically. The meaning of this was given in section 2 above, and mechanisms by which ergodic exploration might be realised in section 3. One way of experimentally testing for ergodic evolution, therefore, is to evaluate whether or not a phenotype of optimal fitness has been found (within other effectively stationary constraints). Unfortunately, the difficulty of ascertaining an optimum fitness in the majority of cases therefore places an obstacle to any direct way of testing whether ergodic exploration of the space has occurred or not.

However, there are some sub-systems whose performance can be evaluated in terms of physics, and judged against constraints that are not contingent on other emergent evolved structures. When this is the case, the phenotype of optimum fitness can be identified in advance from material and physical constraints rather than contingent biological ones. Furthermore, when the history of phenotypes can be tracked through paleontological methods, information can be gleaned on the dynamics of the evolutionary search for optimal fitness. A number of candidate evolved systems of this type are presented by the optics of light-detecting organs. This is because the transmission, focussing, efficiency of detection and diffraction of light are all governed by physical processes of the electromagnetic field and its interaction with matter.

A celebrated example is the invertebrate compound eye. If we assume that the contribution of the insect eye to overall fitness is positively dependent on its angular resolution, or acuity, then the optimal structure of an eye, given its size, becomes a question in physics. First addressed by Barlow in 1952[29], the problem can be reduced to the question of the number (or equivalently the size) of individual optical elements (*ommatidia*) in a compound eye of given radius R. The essential balance to strike is between the *geometrical* optics of the pixilation of an optical field, which limits the angular resolution to (at best) the angular aperture of each ommatidium $\theta = d/R$, and the *physical* optics of diffraction, which limits the resolution of a single ommatidium to $\theta = \lambda/d$, where λ is the wavelength of light detected (see figure 2).
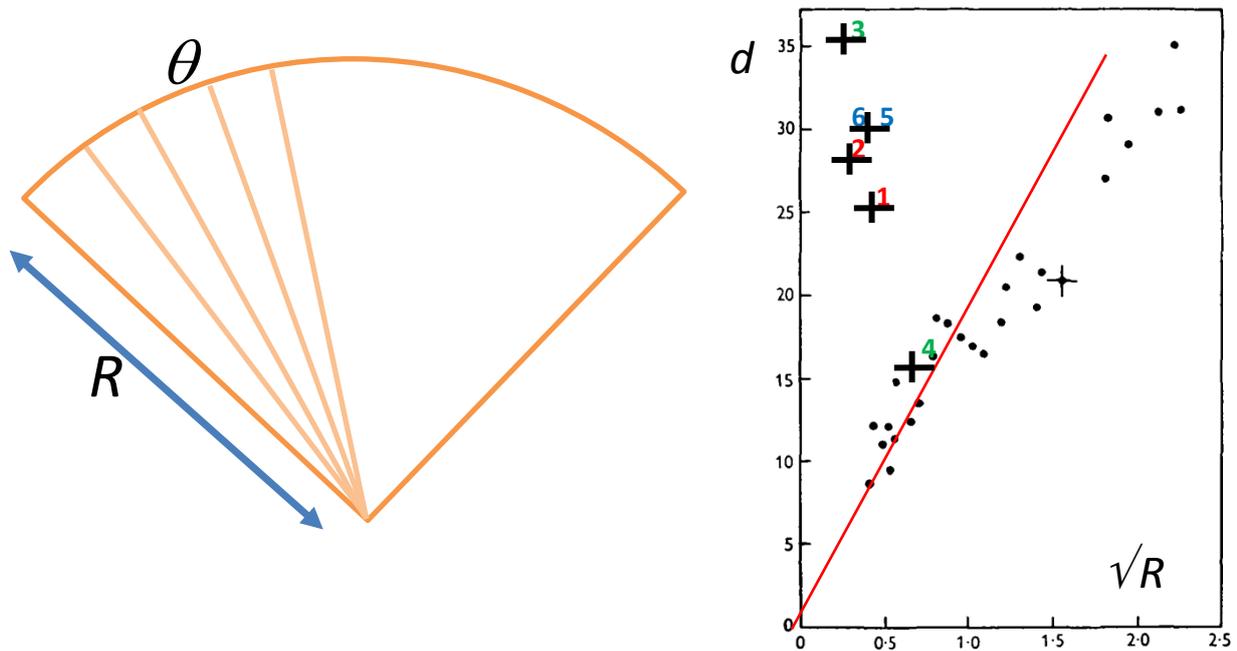
Figure 2 (a) cutaway schematic of a compound eye divided up into ommatidia of angular size θ; (b) and a representative selection of data from pre-historic insects preserved in amber (see table 1), with the optimal line of equation (1) superimposed for a wavelength of 400nm. Units of d are μm, of $\sqrt{R}$, $(mm)^{1/2}$. Error bars for the measurement of micrographs for the 6 archaic cases are given in both coordinates.

Maximising the resolution over the choice of *d* given *R* gives a functional dependence between the two of

$$d_{opt} = \sqrt{\lambda R} \tag{10}$$

This curve is superimposed (as a straight line on the plot in figure 2) together with Barlow's measured results on 27 species of hymenoptera. The curve uses a value of λ=400nm as a reasonable mid-range representative value of the sensitivity of insect photoreceptors[30], but we note that the position of the line is weakly dependent on the precise choice (which in any case represents a spread of about the same order as the spread of data points in Barlow's plot). The conclusion is as remarkable now is it was then, and even more so in the knowledge of evolutionary spaces. Every one of these species has found a size ratio of its ommatidia to the size of the eye itself close to the optically optimal value for that eye radius. We note that the relation (10) does not describe a trivial scaling up of structure, nor a constant optimal ommatidium size. In each case, the species has found how to maximise its optical resolution given its size (eye size tends to scale with overall body size.

This is a phenotype (of a subsystem – the compound eye) which has found an optimum stably over time, so it is very persuasive to identify the optimum resolution (which we know has been found) with an optimum fitness (which is the driver for finding it). Furthermore, from the already-established enormity of the genome space (especially if this phenotype is encoded by several genes) – far too large to explore completely in evolutionary time – then this implies that the genome space has been explored ergodically in the definition of section 2. Without that it is an enormous surprise that $d_{opt} = \sqrt{\lambda R}$ for all the species Barlow measured. Just because we are

13

accustomed to seeing optimally evolved structures does not mean that we should not be astonished by it when we think about what it implies for the exploration of the genome space.

This is clear evidence that the genotype space corresponding to the phenotype of insect eye structure has been explored ergodically in evolutionary time, in the sense that we defined above. Ergodic exploration of genome space is not, as we have seen, exhaustive – indeed it cannot be, but it is sufficient to realise fitness maxima. This example illustrates the evolutionary equivalent, that we have been considering, of the Levinthal Paradox in protein-folding. To understand the detailed structure of the genome space, and its navigation through evolutionary time, that leads to optimal ommatidium spacing, we would need complete knowledge of the protein networks for eye development in each species. This is far from a complete domain of knowledge at present, but the considerable progress that has been made does allow us to make some links to the 'ergodic strategies' of section 2.

Although the vast bulk of the literature on invertebrate compound eye genetics concerns the fly *drosophila melanogaster*[31], the control network for this species' eye development is likely to be of the same order of complexity as in other insects. The insect compound retina is established at an early stage (the 'third instar') of embryonic development[32]. The mosaic of pluripotent cells constituting the retina is differentiated under the passage of a collective linear structure, the 'morphogenetic furrow' from posterior to anterior. As the furrow passes cells, they cluster more tightly. Anterior to the furrow is a region of upregulation of a key signalling protein – a helix-loop-helix transcription factor, Atonal, coded by the gene *ato*. Mutation studies show that *ato* is vital to the regularity and the spacing of the hexagonal array of ommatidia [33]. However, a complex of autoregulation and co-regulation of other proteins has also been implicated in the differentiating and patterning function within the morphogenetic furrow. Posterior to the furrow is a region of equally highly upregulated developmental control protein, Hedgehog (*hh)*, which, along with proteins Notch (*notch*) and Eyeless (*eye*) enhances *ato* expression locally. However, local expression of Notch along with a ligand of Atonal (Dl) and any upregulation of Scabrous (*sca*) *represses ato* in neighbouring cells[34]. The emergent regulatory loop is endowed with spatial structure *via* the spatial diffusion of the Delta (Dl) ligand, Notch and Scabrous proteins. The upregulation of *ato* emerges first as periodic in one dimension along the morphogenetic furrow, then as a 2-dimensional array in the furrow's wake. At the loci of high expression, epithelial cells differentiate into clusters around R8 precursor cells, which pattern the latter full development of the ommatidia, but define their spacing early. More recently, the Epidermal Growth Factor Receptor (EGFR) pathway has been shown to affect R8 spacing and regularity independently of the Notch and Sca pathway[35].

A plausible simplified model for ommatidia patterning, consistent with this known gene network, was proposed recently by Lubensky *et al.*[36]. It successfully represented the development of wild-type patterns as well as some of the mutant pathologies observed. Altogether, a very conservative estimate of the length of the total genome coding for this genetic circuit is 4000 codons. The naïve total genome hypervolume $\Omega = 20^{4000}$ (or even 64^{4000} if we consider all possible codons) A search through this space is unable to account for the repeated optimisation illustrated by figure 2(b).

However, the model of Lubensky *et al*. hypothesises that the 2-dimensional lattice of R8 cells is itself templated by a one-dimensional periodic concentration fluctuation *on its boundary* (the morphogenetic furrow that moves anteriorly across the eye-disk in an early stage of development): the emergence and suppression of *ato* and its concomitant factors serves to propagate a hexagonal array of signalling proteins with the same spacing as the initial periodicity on the morphogenetic furrow (the developmental wave front across invertebrate embryonic eyes has been observed many times). Later fine detail within the ommatidia (including multiple photoreceptors sensitive to different wavelengths *etc.*) appears at a subsequent stage of development. It is likely therefore that

the genetic control of this detailed structure is independent of that which determines the coarse-grained structure of the spacing of the ommatidia.

The genetic control of the initial templating of *ato*-rich periodicity along the morphogenetic furrow is not currently known, but is likely to be much simpler than the entire network required for propagation and downstream fine structure. So this system illustrates beautifully the employment of strategy 3.3 to define a small genetic subspace that can be explored independently of other aspects of the structure, to modify the key variable – ommitidia spacing – in question. Furthermore, within the small region of genome space where the control lies, mutations that fail to explore simple modifications of ommatidium spacing are most likely to disrupt eye structure completely, and cannot be carried through an evolving population – an application of strategy 3.1. All possible compound eyes, independent of the value of spacing chosen, are viable (in the absence of strong competition from an existing optimum) – the percolated criterion of 3.2. Finally, the minimal finite number of mutations ('evolutionary distance') required to attain the Barlow optimum line is not excessive and unlikely to number more than 100 (criterion 3.4)

As a control for Barlow's investigations, and to see if the record of evolution contains any direct evidence for an ergodicity time τ_e, it is instructive to carry out the same exercise for archaic species at a much earlier stage of evolution. Fortunately our example assists here as well, since insects from earlier ages can be preserved in amber to high degrees of anatomical detail. Superimposed on the 'Barlow plot' of figure 2 are 6 examples representative of the literature on amber-preserved insects of various periods from the Cretaceous to the Eocene. This sample of six recent such reported studies was taken from recent literature, filtered only by requiring that the work reported micrographs micrographs at the scale of both the full eye and the ommatidia, both with scale-bars, and that the sample could be dated.   Details are recorded in table 1.

Code number	Reference	Age (MYr)	d (μ) and (residual)	R (μ)
1	[37]	120	25 (+15)	200
2	[38]	120	28 (+20)	80
3	[39]	15-45	37 (+30)	75
4	[40]	15-45	15 (0)	450
5	[41]	40	30 (+22)	220
6	[42]	20	30 (+22)	200

Table 1. Data from amber-preserved insects plotted in figure 2. The age is quoted directly from the published reference in each case. Ommatidium and eye diameters as measured from the published micrographs are given along with the residual ommatidium diameter separating the measured values from the optimum value calculated from equation (10).

It is immediately clear that these earlier species do not typically give optimal ommatidium spacing. Furthermore, they cluster around a very different line on the diagram, suggestive of a trajectory that begins at small R (small eye radius) and large d (large ommatidium size) and that subsequently evolves towards one of the points on the optically optimal curve over timescales of 20-100 MYr. As these species are not on the same line of descent, there is no requirement that their approaches to the Barlow line be well-ordered in time, as indeed they are not. However, the repeated observation of archaic ommatidia away from the current maximal acuity line is very suggestive. Further, properly exhaustive, work on genotype-phenotype mapping within identified lineages will be required to confirm the suggestion. If this example is confirmed as illustrative of the main subject of this paper, the data suggests an ergodicity time τ_e of order 10-100MYr.

Does this seem reasonable given the criteria of section 3 and the final expression for ergodic genome exploration rate (equation 12)? Using current data on effective population size ($N_e \cong 10^6$), rate of deleterious mutations (and so w) and generation time (30 days) for *Drosophila* (recognising

that this can be no more than indicative of a likely order of magnitude for the range of species considered), this range of ergodicity times is consistent with equation (12) for an effective sub-genome of 15 codons. It is not unreasonable to associate a short sequence such as this within one or two proteins for the control of one lengthscale – the one-dimensional lattice of R8 sites on the morphogenetic furrow.

It is possible that the divergence from ideal acuity in most of the archaic species arises not from the lack of ergodicity and consequent failure to optimise, but on other environmental constraints. These (such as sensitivity in low light levels) may mean that these eyes are optimised. Only detailed lineage analyses will tell. As one extreme limiting data point, however, it is worth noting that the earliest compound eye (or at least a much earlier one that those of figure 2 and table 1) is found in a lower Cambrian trilobite from 400MYr ago[43]. These, structures, strongly reminiscent of today's insect ommaditia are of much larger dimension ($d>200\mu$), and would be well outside the top edge of the graph in figure 2. It is not unreasonable that early compound eyes would have developed with a coarse-grained design, whose development became successively finer under evolutionary pressure.

5. Discussion and Conclusion

The nature of evolutionary spaces is very high-dimensional, but structured in a way that allows a genetic equivalent of ergodic exploration to be realistic within a timespan of tens of Myr. This is highly suggestive of the property of ergodicity in dynamical systems theory and fundamental statistical mechanics, especially since there is already a large body of literature applying notions and methodologies from statistical mechanics to models of evolution.

However, the approach to ergodicity, in both its definition and application, needs to be different from that in dynamical systems. There are no conserved quantities, nor deterministic equations of motion. However, the fundamental notion of representative exploration of state space, in which a dynamical trajectory visits arbitrarily closely any given state point, does survive. There will be qualifications – for example notion of weight of points must be employed (there are unviable regions of genome space that do not permit any trajectory to pass close to them).

In evolution, both sexual and asexual, we have examined three strategies that together generate an ergodic exploration of the genome: confinement, percolation, hierarchical dimensional partitioning. These work together with a fourth property of genome space, which limits the generalised distance of two genomes to the order of their own length, to deliver the essential dynamic and geometrical structures that make possible the convergence to optimal fitness in sufficiently compact subsystems. Although we have suggested mathematical routes to explore, which capture these four properties, it remains to quantify the hypothesis of ergodic genome exploration. If this is possible then it will deliver a predictive ability to determine which phenotypes have been potentially optimised through ergodic exploration in evolutionary time.

A summary of the mapping from statistical mechanics and ergodic theory to evolutionary dynamics might usefully be tabulated. In some cases these are closely mathematically related, in others there are properties in evolution that replace those in dynamical systems.

Statistical Mechanics	Evolutionary Dynamics
High dimensional State space (phase space)	High dimensional Genome space
Macrostate	Phenotype
Equations of motion	Mutation of the genome
Ergodicity	Representative exploration of the genotype-phenotype map
Low-dimensional embedded trajectories	Pathway confinement
Conserved quantities,	Percolation and branching,

Exponential Divergence from initial conditions	Hierarchical genome subspaces
Free energy optimisation	Free fitness optimisation

Table 2. Corresponding concepts between ergodic theory and statistical mechanics (left hand column) and evolutionary dynamics (right hand column).

Furthermore, when optima are physically constrained, it suggests the possibility in some special cases of *a priori* prediction of cases of convergent evolution. This is that when (i) a system possesses ergodic evolutionary dynamics and (ii) its optimum fitness can be determined from physical constraints, then it is possible to predict the (convergent) outcome. So, taking the example of section 4, if we had identified (i) and (ii) in advance, it would be possible to predict that the 'Barlow relationship' would arise for ommatidia spacing.

This specific case of insect compound eyes promises to be a productive first test-bed for these ideas. But even here there is much to do. The gene control network for ommatidia spacing needs to be isolated. Then as far as possible, past mutations within that network need to be mapped on to the convergence towards the 'Barlow' curve of optimal acuity in lines of descent. Finally a specific model of the ergodic structure of this genomic subspace needs to be constructed from the processes outlined in section 3.

The possible application of these ideas to other convergent systems for which fitness optima arise from physical constraints is open to exploration. Upstream organs of perception, since these respond to physical constraints such as those of physical optics, would suggest themselves as promising candidates.

References

[1] H. P. de Vladar and N. H. Barton "The contribution of statistical physics to evolutionary biology", *Trends in Ecology and Evolution*, **26**, 424-432 (2011)

[2] Fisher, R.A. (1922) 'On the dominance ratio' *Proc. R. Soc. Edinb.* **42**, 321–341

[3] Wright, S. (1931) Evolution in Mendelian populations. *Genetics* **16**, 97–159

[4] Barton, N.H. and Coe, J. (2009) On the application of statistical physics to evolutionary biology. *J. Theor. Biol.* **259**, 317–324

[5] Y Iwasa, 'Free fitness that always increases in evolution', *J. Theo. Biol.*, **135**, 265–281 (1988)

[6] Sella G, Hirsh AE The application of statistical physics to evolutionary biology. *Proc Natl Acad Sci USA* **102**, 9541–9546 (2005)

[7] Berg J, Willmann S, Lässig M Adaptive evolution of transcription factor binding

sites. *BMC Evol Biol* **4:**42. (2004)

[8] Barton, N.H. and de Vladar, H.P., 'Statistical mechanics and the evolution of polygenic quantitative traits', *Genetics*, **181**, 997–1011 (2009)

[9] L. Boltzmann. See part 2 of *Vorlesungen über Gastheorie*. Leipzig: J A Barth 1898

[10] E. Artin, "Ein mechanisches System mit quasi-ergodischen Bahnen", *Abh. Math. Sem. d. Hamburgischen Universität*, **3** 170-175 (1924)

[11] Ya. G. Sinai, "Dynamical Systems with Elastic Reflections", *Russian Mathematical Surveys*, **25**, (1970) 137–191.

[12] Chan, H. S. & Dill, K. A. 1991 "Sequence space soup" of proteins and copolymers. *J. Chem. Phys.* **95**, 3775–3787.

[13] D. T.F Dryden, A. R Thomson and J. H White, 'How much of protein sequence space has been explored by life on earth?' *J. R. Soc. Interface* **5**, 953-956(2008)

[14] S.J. Gould, *Wonderful Life*, Vintage Books London (2000)

[15] Conway Morris, S. (2004) *Life's solution: inevitable humans in a lonely universe*. Cambridge, UK: Cambridge University Press.

[16] David L. Stern, and Virginie Orgogozo (2008) *"The Loci Of Evolution: How Predictable Is Genetic Evolution?" Evolution* **62-9***: 2155–2177*

[17] Gavrilets, S. (1997) Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol*. **12**, 307–312

[18] H. Meyer-Ortmanns and A. Klümper Eds, *Field theoretical tools in for Polymer and Particle Physics*, Springer Lecture Notes in Physics **508** (1998)

[19] L. Loewe and B. Charlesworth, 'Inferring the distribution of mutational effects on fitness in Drosophila', *Biol. Letts*., 2, 426–430 (2006)

[20] D. Stauffer and A. Ahrony, *Introduction to Percolation Theory*, Taylor and Francis London (2010)

[21] J. Gravner, D. Pitman and S Gavrilets, Percolation on fitness landscapes: effects of correlation, phenotype and incompatibilities, *J. Theo. Biol.*, **248**, 627-645 (2007)

[22] Adam, G., and M. Delbrück, 'Reduction of dimensionality in biologicaldiffusion processes. *In* Structural Chemistry and Molecular Biology', A. Rich and N. Davidson, *eds* W. H. Freeman & Co., New York. 198-215 (1968)

[23] E.I. Shakhnovich, R.A. Broglia and G. Tiana Eds., *Protein Folding Evolution and Design*, IOS Press Amsterdam (2000)

[24] Rimratchada, Supreecha; McLeish, Tom C. B.; Radford, Sheena E. and Paci E., "The Role of High-Dimensional Diffusive Search, Stabilization, and Frustration in Protein Folding", *Biophys. J.*, **106**, 1729-1740 (2014)

[25] K.A. Dill, S.B. Ozkan, M.S. Shall and T.R. Weikl, 'The Protein folding Problem', *Ann. Rev. Biophys.,* **37**, 289-316 (2008)

[26] Phillips, P.C. 'Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems', *Nature reviews. Genetics* **9**, 855–67 (2008)

[27] Stearns, F. W., 'One Hundred Years of Pleiotropy: A Retrospective', *Genetics* **186**, 767-773 (2010).

[28] Cathy Haag-Liautard *et al.*, 'Direct estimation of per nucleotide and genomic deleterious mutation rates in Drosophila', *Nature*, 445, 82-85 (2007)

[29] H. Barlow, 'The Size of Ommatdia in Apposition Eyes', *J Exp Biol*, **29**: 667–674 (1952)

[30] R. Menzel, D. F. Ventura, H. Hertel, J. M. de Souza, and U. Greggers, 'Spectral sensitivity of photoreceptors in insect compound eyes: Comparison of species and methods', *Journal of Comparative Physiology A*, **158**, 165-177 (1986)

[31] M. Tsachaki and S, G. Sprecher, 'Genetic and Developmental Mechanisms Underlying the Formation of the Drosophila Compound Eye' *Developmental Dynamics* **241**, 40–56 (2012)

[32] L. R. Squire, *Fundamental Neuroscience*, Academic Press (2013)

[33] N. M. White and A. P. Jarman, 'Drosophila Atonal controls photoreceptor R8-specific properties and modulates both Receptor Tyrosine Kinase and Hedgehog signalling', *Development*, **127**, 1681-1689 (2000)

[34] J. P. Kumar and K. Moses, 'EGF Receptor and Notch Signaling Act Upstream of Eyeless/Pax6 to Control Eye Specification', *Cell*, **104**, 687–697 (2001)

[35] B. J. Frankfort and G. Mardon, 'R8 development in the Drosophila eye: a paradigm for neural selection and differentiation', *Development*, **129**, 1295-1306 (2002).

[36] D.K. Lubensky, M.W. Pennington, B.I. Shraiman and N.E. Baker, 'A dynamical model of ommatidial crystal formation', *PNAS,* **108**, 11145-11150 (2011)

[37] N. Barling, D. M. Martill, S, W. Heads, F. Gallien, 'High fidelity preservation of fossil insects from the Crato Formation(Lower Cretaceous) of Brazil', *Cretaceous Research* (2014) in press

[38] J. Koteja and D. Azar, 'Scale insects from Lower Cretaceous amber of Lebanon (Hemiptera: Sternorrhyncha: Coccinea)', *ALAVESIA*, **2**: 133-167 (2008)

[39] M. Kogan and G. Poinar Jr, 'New fossil Stylops (Strepsiptera: Stylopidae) from Dominican amber', *Neotrop. entomol.* **39** no.2 (2010)

[40] G. Tanaka, A. R. Parker, D. J. Siveter, H. Maeda and M.Furutani, 'An exceptionally well-preserved Eocene dolichopodid fly eye: function and evolutionary significance', *Proc Roy Soc B*, **276**, 1015-1019 (2009)

[41] J. Batelka, 'Olemehlia krali, a new genus and species of Ripidiinae from Baltic amber (Coleoptera: Ripiphoridae)', *Ann. soc. entomol. Fr. (n.s.)*, **47**, 361-364 (2011)

[42] J. Batelka, M. S. Engel, Z. H. Falin and J. Prokop, 'Two new ripidiine species in Dominican amber with evidence of aggregative behaviour of males "frozen" in the fossil record Coleoptera: Ripiphoridae)', *Eur. J. Entomol.* **108**: 275–286, (2011)

[43] B. Schoenemann and E, N. K. Clarkson, 'Discovery of some 400 million year-old sensory structures in the compound eyes of trilobites', *Scientific Reports,* **3**, 1429 (2013)