Neutral Spaces and Topological Explanations in Evolutionary Biology: Lessons from Some Landscapes and Mappings

Philippe Huneman*†

I consider recent uses of the notion of neutrality in evolutionary biology and ecology, questioning their relevance to the kind of explanation recently labeled 'topological explanation'. Focusing on fitness landscapes and genotype-phenotype maps, I explore the explanatory uses of neutral subspaces, as modeled in two perspectives: hyperdimensional fitness landscapes and RNA sequence-structure maps. I argue that topological properties of such spaces account for features of evolutionary systems: respectively, capacity for adaptive evolution toward global optima and mutational robustness of genotypes. Thus many models appealing to "neutral" manifolds provide topological alternatives to hypothetical mechanisms.

1. Introduction. In this article, I explore the connections between a concept that emerged in evolutionary biology and ecology some decades ago and is now pervasive, namely, 'neutrality', and a kind of explanation to which some philosophers recently paid attention, namely, 'topological explanation'. I argue that some instances of neutrality in evolution and ecology are the locus of deployment of characteristic topological explanations.

*To contact the author, please write to IHPST, CNRS/Université Paris 1 Sorbonne, 13 rue du Four 75006 Paris, France; e-mail: philippe.huneman@gmail.com.

I will focus on two modeling ideas that arose in evolutionary and ecological modeling since 2000, concerning two classical tools in evolutionary theorizing, namely, fitness landscapes and genotype-phenotype maps. Landscapes and maps are obviously likely to be interpreted in topological terms. But in recent years, researchers have started to consider how neutrality (in the evolutionary sense defined below) can affect them. I will argue that new kinds of explanatory practices are involved here.

The first section of the article summarizes views on topological explanations. The second section presents 'neutrality' in evolution and ecology. The third section explains the idea of 'neutral tunnels' in fitness landscapes and makes sense of the explanations that go on there; the fourth considers 'neutral networks' in genotype-phenotype maps and focuses on the explanatory practices involved. The last section draws consequences of those explorations by arguing that the notion of 'neutrality' as it developed in the last decade naturally appeals to a set of explanatory practices that essentially involve topological explanations, and that in turn the explanatory repertoire of topological explanations should be reconceived.

2. Structural and Topological Explanations. Philosophy of science in the 2000s has been widely interested in an account of scientific explanations that takes them to be the unraveling of mechanisms (e.g., Craver and Darden 2013), namely, a set of entities endowed with proper activities, organized in a specific way, and likely to yield the explanandum. Yet recently some philosophers pointed out that there exist explanations unlikely to consist in unraveling mechanisms. For instance, they considered optimality explanations in behavioral ecology (Rice 2012) or statistical explanations (Lange 2013) based on pure laws of statistics. 'Topological explanations' are similar in this respect: they explain the explananda by appealing to topological or graph-theoretical properties of a space attached to the system under focus and then showing how it is entailed from them (Huneman 2010; Jones 2014).

In all those explanations, it seems that the precise mechanism itself is not explanatorily relevant; similarly, knowing the exact trajectory of the system indeed does not provide the explanation; another characteristic is their genericity. Generally, in these explanations a formal property plays a crucial role in explaining the explanandum. Granted, in all explanations, mathematics can be found; but what is proper to the class of explanations I just cited is that mathematics is not playing a representational role (for instance, functions represent the effects of the activities of entities in a mechanism) but an explanatory role. Of course, the models are still supposed to correctly represent a target system, but the 'representation question', about whether a model represents its target, is only one epistemic question, likely to be raised about any kind of model and explanation. Here I focus on another question, namely, "how does the explanation account for what is supposed to be explained?" In those explanations deemed nonmechanistic, mathematical properties, be they

topological, statistical, or other, are precisely doing the accounting job; subsequently, they play an explanatory and not representational role.

Huneman (2018) proposed on this basis a view of those explanations (hereafter labeled 'structural explanations'), intending to make sense of the way mathematical properties play such an explanatory role. In this view, on which I will rely here,

> *P* is a mathematical property playing an explanatory role regarding an explanandum in a system *S* iff *P* is a mathematical proposition—holding of a space of mathematical entities—that constrains any mathematical representations of the mechanisms or trajectories of *S*, therefore contriving it to display a set of generic features.

Accordingly, *topological explanations* are the structural explanations in which the explanatory role is played by a mathematical fact pertaining to topology or graph theory. They explain a feature such as stability or robustness of a system by deriving the consequences of a topological property of a space proper to the system, such as being a scale-free network. This contrasts with an explanation that would unravel, for instance, a feedback process restoring a state of the system after perturbation: while the former are topological, the latter is indeed clearly a mechanistic explanation.

3. Neutrality, Evolutionary and Ecological. What neutrality means in evolutionary biology has been put to the fore of the discipline by the 'neutralist theory' of evolution, developed by Kimura, among others, in the 1960s and 1970s. 'Neutral' here is about mutations that are such that they do not affect fitness: they are therefore neutral regarding natural selection. Wright (1931) emphasized that selective processes in small populations would undergo 'genetic drift', namely, a sort of sampling error due to the small size of the populations. Drift is a notoriously difficult concept (Plutynski 2007): some see it rather as a process (e.g., Millstein 2002) and others as an outcome of evolution in small populations (Gillespie 2004); some define it by including many random processes in it (gamete sampling, founder's effect, etc.) while others view it rather in mathematical terms as a deviation from frequency expectations based on fitness (Matthen and Ariew 2002). Notwithstanding the concept chosen, since an evolution driven by drift does not predictively yield results according to fitness values (w) one can distinguish between drift (type 1) as the case in which two alleles distinct in fitness ($w < w'$) compete but the lower fitness allele goes to fixation, and drift (type 2) as what happens when two equal fitness alleles ($w = w'$) compete, and one goes extinct. Many concepts of drift consider them together (notice that type 2 is the limit of type 1 when $w \rightarrow w'$). Kimura's (1983) neutralist theory focuses on the drift type 2, the only one about neutrality, and argues that much of the composition of the genome at the nucleotidic level is shaped

by this drift. The abundance of DNA changes that are neutral is an argument for the claim that chance plays a major role at the nucleotidic level (Crow 1987, 12).

Ecologists more recently designed a neutralist theory that is symmetrical to the neutralist theory in evolution. This 'unified neutral theory of ecology', systematically elaborated by Hubbell (2001) around the assumption of fitness equivalence between individuals in different species, objects to the leading Darwinian selectionist theory in ecology in the same way as Kimura objected to the selectionist viewpoint and has been very predictive of many biodiversity patterns (coral reefs, tropical forests; McGill, Maurer, and Weiser 2006).

The common key feature is that in any of those neutral models, the biological difference—between species, between alleles—does not make a difference to the explanandum; this is the whole epistemological point about having an equivalence hypothesis. Therefore, it is difficult to think about explanations given by a neutral theory as causal explanations, to the extent that 'causation' would be understood in terms of difference making (Lewis 1973; Woodward 2003). So there is philosophical room here for explanations that would not fit causal or mechanicist views of explanations. I now turn to some theoretical uses of neutrality in evolutionary biology.

4. Fitness Landscapes and Neutral Tunnels. Fitness landscape is a pervasive tool in evolutionary biology (fig. 1). It is a surface in a hyperdimensional space, intended to capture the fitness structure of a genotype pool and its potential for evolution. Designed initially by Sewall Wright (1931), it can receive two interpretations: either the axes can be the alleles themselves, and then the value on an axis is the frequency of this allele in the population, or they can be the genes on a locus, and then the values on an axis are the alleles. The vertical axis is always the fitness. In the former case the points are populations, defined by a set of allele frequencies; in the latter case they are organisms, so that the population is a cloud of points. Those two interpretations are not orthogonal, and indeed there exists formal correspondences between both of them.

Since their inception fitness landscapes—sometimes termed 'adaptive landscapes'—underwent a controversial fate. Fisher resisted them, yet one of the main assumptions of fitness landscapes is that natural selection always drives populations or organisms (points or clouds of points) toward higher fitness states. As a consequence, Wright had to appeal to Fisher's 'fundamental theorem of natural selection', even though this theorem has received various conflicting interpretations over the years.[1] Fisher, however, would

1. The most recent consensus (after Price [1972]) is that the intergeneration variation in fitness due to natural selection is equal to additive genetic variance, which is by definition always positive.
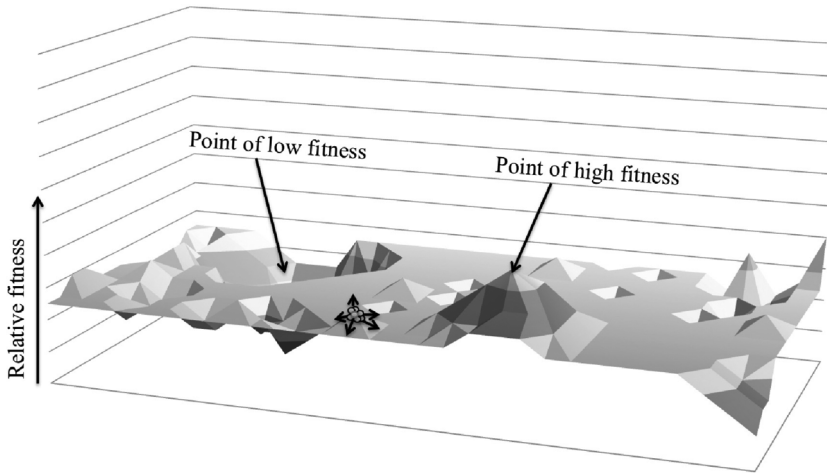
Figure 1. A fitness landscape, with global and local optima. Color version available as an online enhancement.

object to this use of his fundamental theorem: for him Wright used it as a physicist, in the same sense as potential wells, whereas he thought this assimilation made no sense (see Frank 2012).

Those formal devices can address a specific problem, namely, epistatic genetic interactions. In the absence of interactions, fitness of alleles can be treated additively (each allele contributing additively to the fitness value of the genotype). So (writing $W(X)$ for 'the fitness of X') given two loci and two alleles for each (x, X, y, Y), if $W(xy) < W(XY)$, then we can assume that $W(xy) < W(xY)$ and $W(xy) < W(Xy)$, and then $W(xY) < W(XY)$ and $W(Xy) < W(XY)$. So, in order to increase $W(xy)$, a system can independently go through Xy and then XY or through xY and then XY: in any case, we follow a path of increasing fitness. Thus selection is the only process needed to go from xy to XY, whatever the trajectory (i.e., which allele gets fixed first). This process can indifferently go through changing x to X first or y to Y first.

However, what if there is no additivity, but the fitness value of xy does not increase similarly when one changes x to X or y to Y? Suppose instead that changing only one allele (x or y) decreases the fitness value of the diploid genotype. This is 'epistatic' interaction, and its signature is $W(xy) < W(XY)$ but $W(xY) < W(xy)$ and $W(Xy) < W(XY)$. Since going from xy to XY implies changing successively each of the two alleles, it means that the fitness value of the genotype has to decrease at some point and then increase later on. Then it is not conceivable that natural selection—supposedly always increasing fitness—is the only process going on.

Put in terms of fitness landscape, the problem is obvious: to go from xy to XY, the system has to drop in fitness and then reclimb a fitness hill. We

are facing a problem of being stuck on a local optimum: xy is locally very good so selection cannot improve it. There exists another optimum, better in fitness, namely XY, but natural selection is unable to reach it. That is where genetic drift comes into play.

Wright hypothesized that drift is the process that allows the population to decrease in fitness and then reach the genotype xY (or Xy), lower in fitness than xy; once this is done, selection is now able to push the population to a better fitness optimum, namely XY. The same can go on until the population reaches the genuine global optimum. Thus what Wright proposed here under the name of shifting balance theory (SBT) is a mechanism to explain how populations can reach adaptation even in a case of pervasive epistasis. Of course it assumes that biological populations are small, so that drift plays an important role in evolution. It includes three stages: (1) genetic drift, which takes the population down, away from the local fitness peak; (2) selection, which makes it climb the next fitness peak; and (3) migration of high-fitness individuals.

This hypothetical mechanism has been criticized (Coyne, Barton, and Turelli 1997): empirically, few instances have been found, and mathematically it seems that its conditions are strong; the conditions for stage 1 (drift) are different from the conditions for stages 2 and 3, which makes the overall requisites for the mechanism very demanding.

This is one main reason for the doubt several theorists entertain regarding fitness landscapes. However, a major and obvious critique is the fact that the landscape itself is not immutable. Since the fitness of alleles depends on the composition of the population in cases of frequency dependence, across time the landscape will change as a consequence of selection within itself. Yet the moving landscape generally remains in between certain theoretically determinable limits; hence for an evolving population the landscape can still be understood as a characteristic feature likely to capture its evolution.

More recently, however, Gavrilets (1997, 1999) switched the focus to a major feature of fitness landscapes: they are hyperdimensional. Therefore, the intuitions based on three-dimensional space actually do not apply to those spaces. But precisely, most of the reflections on them were based on those intuitions, even if Wright was careful enough to mention that those landscapes are high-dimensional. And most importantly, our intuitions about local and global optima in Euclidean three-dimensional space do not carry over straightforwardly for hyperdimensional spaces.

In three dimensions indeed, to go from xy to XY (such that $W(xy) > W(xY)$ but $W(xy) < W(XY)$), one cannot but drop in fitness along axis W. However, in higher dimensions there exist ways to go from xy to XY without losing fitness, because on the landscape there is a hypersurface of constant fitness on which the point moves, in a way that its values over other dimensions (but not the axis W) change. Then by drifting on this manifold (in the mathematical sense), which Gavrilets called a 'neutral tunnel', the system can reach

another neighborhood whose fitness maximum is higher (fig. 2)—even if strictly speaking no hills and valleys exist in hyperdimensional spaces. This is impossible in three dimensions because, in order to remain on the landscape and move on axes x and y away from a local fitness peak W_0, the only axis on which the system can move is the W axis.

In hyperdimensional space, landscapes may thus appear as different (Gavrilets 1997, 1999). The theoretical upshot is crucial: no drift stage in the sense of type 1 drift ($w < w'$) is needed to explore a valley and then find another fitness peak; the exploration can take place only within the neutral tunnel. Hence appealing to the three stages of the SBT mechanism hypothesized by Wright (and whose feasibility has been questioned since Coyne et al. [1997]) is not necessary to solve the problem of adaptive evolution in highly epistatic genetic systems, which raises the general issue of how to avoid getting stuck on a local fitness peak—an issue that must be tackled in order to make sense of adaptive evolution in general.

Importantly, the sense of 'drift' here therefore differs from the drift stage in SBT: in SBT it is type 1 drift, about alleles or genotypes of distinct fitness, and concerns the fixation of the low-fitness one (against what fitness would predict); in the case of our neutral tunnels, it is type 2 drift (as in Kimura's
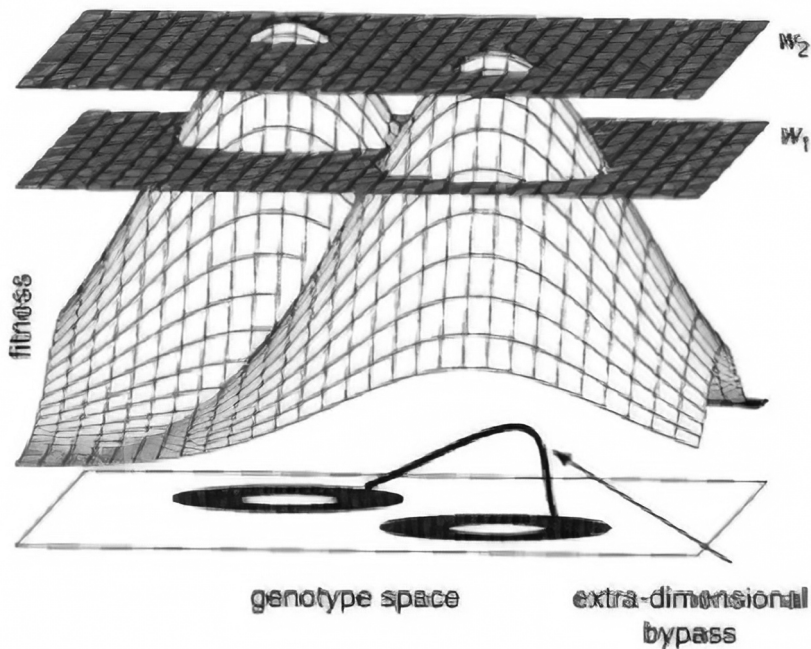

Figure 2. Figuring in three dimensions the between-hills neutral tunnels (as planes) in fitness landscapes (after Gavrilets [1997]).

neutralist theory), namely, a trajectory within a set of equal fitness geno-types.

The possibility of a pathway between two distinct local optima in fitness (W_1, W_2, and $W_2 > W_1$) relies on the existence of a submanifold—the 'neutral tunnel'—that exists between the two neighborhoods of those local optima, and that is such that W can remain constant in this submanifold, while the values on all other coordinates change. Hence the mathematical fact that such a submanifold exists on a fitness landscape constrains any trajectory of a population from an initial neighborhood with fitness maximum W_1 'within' the tunnel to access a neighborhood where a higher fitness optimum W_2 is met and thus explains the possibility of adaptive evolution; or at least it explains the adaptive evolution from a local fitness optimum W_1 to a higher local fitness optimum W_2, the whole adaptive evolution being a series of such moves.

In turn, the existence of such a submanifold relies on a mathematical fact characterizing spaces of dimensions higher than three. In a hyperdimen-sional space of dimension $n + 1$, a manifold is n-dimensional. Thus with $n > 3$, the manifold constituted by a fitness landscape will be of dimension at least three. It is a fact that tunnels cannot exist in dimensions lower than three since they are at least three-dimensional; in contrast, in a manifold in an n-dimensional space, there may exist a tunnel, namely a submanifold, in which W can remain constant and the other coordinates vary.

Many m-dimensional manifolds with $m \geq 3$ indeed have tunnels, which are defined by the holes that they 'surround'. Manifolds are classified by their 'genus', which is the number of holes (being therefore topologically equivalent). Therefore, m-dimensional manifolds of genus $g \geq 1$ feature tun-nels (fig. 3), some of them being likely to provide fitness-constant pathways. The fact of tunnels, proper to those manifolds of such a genus, accounts for the possibility of adaptive evolution in systems whose fitness landscapes are such manifolds.

Therefore, mathematical facts about hyperdimensional topological spaces make possible the existence of neutral tunnels, which in turn accounts for the possibility of type 2 drift underlying adaptive evolution in case of pervasive epistasis (unlike what goes on in two-dimensional landscapes, where only SBT mechanisms can explain such adaptive evolution). To this extent, topo-logical properties of dimensionality and genus value ground the account of the possibility of adaptive evolution (in cases of systems with a high degree of epistasis); hence, topological explanations provide an alternative to the SBT mechanism explanation.

5. Neutral Networks and Genotype-Phenotype Maps. Genotype-phenotype (GP) maps were introduced in biology once the complexity of the relation between genotypes and phenotypes had been acknowledged
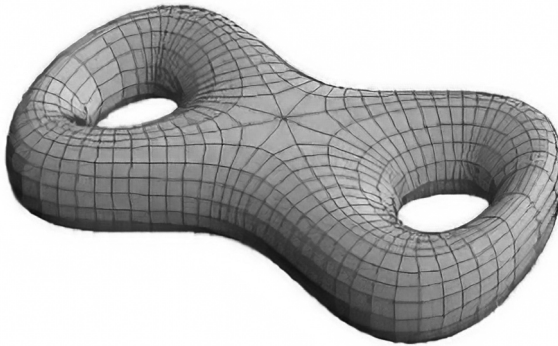
Figure 3. A manifold of genus 2: two holes, which are surrounded by submanifolds, called 'tunnels' in some fitness landscape literature. In such a tunnel may exist a pathway along which the *W* coordinate remains constant. Color version available as an online enhancement.

(e.g., Lewontin 1974). We have known for many decades now that this relation is not a one-to-one. Genotypes impinge on many phenotypic traits, and reciprocally—precisely through epistasis—several genes or genotypes affect the same phenotype. The set of those mappings between genotypes and phenotypes sets the stage for possible evolution. For instance, a phenotypic trait is likely to emerge only if all the genotypes that condition it are available in the population; and inversely, the fixation of a phenotypic trait constrains all traits whose phenotypes are affected by the genotypes underpinning this trait.

The space of genes or genotypes and the space of phenotypes are therefore functionally related.[2] The pattern of this relation (or mapping) determines the space into which evolution, and especially adaptive evolution, is possible. There are many kinds of GP maps.[3] Investigating such maps illuminates the constraints a genetic makeup of a population superimposes on any evolution by natural selection. Some maps elaborated in the 2000s were intended to address robustness and adaptive evolution.

Robustness can be defined as a capacity to keep some parameters stable in the face of changing variables (a context-dependent concept); it is ubiquitous in biology. Cell machinery and DNA (e.g., translation and transduction in the nucleus) can perform their activities through lots of thermodynamical noise; organisms keep constant physiological parameters through homeostatic mechanisms studied by physiologists; and of course embryological

2. 'Function' is taken in the mathematical sense, of course.
3. See, e.g., Winther (2018) for a study of those maps, their history, and their philosophical underpinnings.

development keeps producing organisms specifically like their parents even if it is disturbed, as experimental embryology has established it.

Two meanings of robustness are indeed at stake: environmental robustness, which measures robustness with respect to the variations of the environment, and mutational robustness, which is about how much perturbations on genes ultimately affect the system. 'Canalization', first represented by Waddington as valleys in an 'epigenetic landscape', is of the latter kind: it is about the constancy of the phenotypic output of development relative to tweaking the genes that support development (by mutation, mostly).

However, did robustness appear first as environmental robustness? Or was mutational robustness the target of selection, and then, once robustness evolved it buffered the system against environmental noise too? The question of how mutational robustness is possible has been addressed in terms of GP maps by Stadler, Fontana, and colleagues in the 2000s (e.g., Van Nimwegen, Crutchfield, and Huynen 1999; Stadler et al. 2001). The model they use first considers RNA sequences, understood as genotypes, and their relation with phenotype and fitness. The phenotype is the functional protein, namely, the folded protein; but given that proteins do interact by receptors on their surface, several proteins having a different sequence may feature the same functional profile and therefore the same fitness.

The model therefore considers the network made up of the possible genotypes (each genotype is a node); the distance between two genotypes is defined by the amount of mutations by which they differ (fig. 4). Then the space of genotypes is mapped onto phenotypes and the fitness. Several genotypes are likely to have the same fitness, so Stadler, Fontana, and colleagues define what they call "neutral networks," namely, subsets of the network of genotypes in which all genotypes have the same fitness and such that all genotypes in one of such subset are neighbors; namely, they are separated by one mutation step.

Studying genotypes by pointing out their position within the neutral network clearly gives rise to explanations based on graph-theoretical properties, that is, topological explanations. In effect, neutral networks may be convex subgraphs, namely, subgraphs in which all shortest paths between two vertices are included. Thus, let us consider such a convex subgraph S with a common fitness W and two vertices N_1 and N_2 outside S; their fitnesses W_1 and W_2 differ from the common fitness W of the neutral network, and we conventionally assume the order $W_1 < W < W_2$. Let us take N_1 and N_2 as related to vertices in the subgraph S by only one edge. Then S has the property that the shortest path between those N_1 and N_2 will be included in the subgraph. The reason is that N_1 has a neighbor N and N_2 a neighbor N' such that N and N' are in S, and the shortest path between them is therefore in S. Thus there will be at least one path that links N_1 and N_2 and remains included in the neutral network S. The system therefore is constrained to fea-

Figure 4. Network of genotypes, each node being an RNA sequence. An edge between two nodes represents a distance of one mutation. Nodes of the same color have the same fitness. Black subnetworks are neutral networks.

ture at least one path between N_1 and N_2 that goes from N_1 only through same-fitness genotypes (in S) until it meets the first higher-fitness genotype N_2. Thus, the system should include pathways of adaptive evolution through type 2 drift.

Thus, the mathematical property of convexity of the subgraph 'neutral network' S allows for possible drift to explore the genotype space without losing fitness. Then the position of a focal genotype G strictly within the neutral network implies that there exist several drift pathways going through G—pathways that link N_1 to N_2. Moreover, the larger the neutral network, the higher the amount of such possible drift pathways going through G—and then of drift pathways in general—that allow adaptive evolution. The existence of a neutral network S and its convexity finally account for the possibility of an exploration of the genotype space by drift and a possibly

adaptive evolution when such a pathway leaves the neutral network S at a higher fitness node N_2.

In the same way, those graph-theoretical facts are thereby what accounts for the fact that mutations do not harm the focal genotype (since mutations within a convex neutral network S do not decrease fitness), hence making it mutationally robust. Such an explanation contrasts with possible explanations of robustness that appeal to repair mechanisms, such as sexual reproduction, which allows for repairing damages on a strand of DNA by relying on the second strand as a template.

6. Neutrality as a Locus for Topological Explanations. The neutral network in GP maps and the neutral tunnels in fitness landscapes both account for features of evolutionary systems by considering topologies of abstract spaces attached to them. The GP maps address primarily the question of robustness and then of evolvability; the fitness landscapes address the question of adaptive evolution. While the landscape does not concern phenotypes, but only fitness, the GP map integrates fitness but focuses on the GP relation. They are different mathematical takes on a space of genotypes with a given fitness function.[4] The tunnel and the subnetwork are two kinds of 'manifolds' in a general topology ('topological graph theory' *sensu* Gross and Tucker [1987]).[5] Their properties account for evolutionary features by constraining the exploration of the space by genotypes.

Intuitively, however, the GP map and the fitness landscape are not unrelated. They determine together the formal properties of a general space of genotypes that define the possible evolutionary trajectories before any mechanism of selection (or other) comes into play. A neutral subnetwork in a GP map will correspond to a neutral tunnel in a fitness landscape since, being a 'manifold' of genotypes with equal fitness, once mapped onto the landscape it will become a submanifold of equal fitness value in the landscape, that is, a neutral tunnel. But can we formally unify those spaces?

Both constitute the topology that conditions drift; yet they rely on different informations: no set of fitness values is given in GP maps (just the fitness equivalence relation), and no phenotypes are given in landscapes. One therefore cannot infer the fitness landscape from the GP map or inversely. Hence they are not dual spaces simply connected by an isomorphism, even though they are related because they could be understood as two perspectives on the same formal space of possible evolution of a completely determined set of genotypes defined by all phenotypic and fitness information. They give rise to possible explanations that rely on two aspects of topology: graph theory

4. The epigenetic landscape, in turn, focuses on the GP relation only.

5. From here on I use 'manifold' in a very large sense to cover both kinds of spaces, in accordance with the general meaning given to 'topology' in the phrase 'topological explanation'.

(for maps) and topology *stricto sensu* (for landscapes). Together they feature how the introduction of topological views in evolutionary biology expands our explanatory capacities of evolution in complex systems by better capturing some of the evolutionary potential of neutrality.

Considering them together shows the general connection between neutrality and topological explanations: these explanations appealing to neutral manifolds are not selectionist explanations, for neutral manifolds are defined as subspaces where the biological difference makes no difference (in fitness) and are explanatory precisely in virtue of that. The mathematical structure of the overall space and subspaces—and not the difference between genotypes (and their effects)—plays the accounting role in the explanation via theorems (regarding convexity in subgraphs or existence of tunnels in hyperdimensional topologies) about which possible pathways can occur. In both explanations indeed, stochastic change in genotypes—type 2 drift ($w = w'$)—is constrained in such a way that it can yield by itself actual adaptive evolution or a potential for adaptive evolution. Subsequently, both neutral manifolds provide explanatory alternatives to mechanisms: the neutral tunnel avoids appealing to SBT, which is epistemically costly because very demanding; the neutral network differs from explanations involving repair mechanisms. In a nutshell, in those cases neutrality thereby makes topology relevant to explain adaptive evolution.

Topological explanations have been mostly considered for their intrinsic potential to explain robustness, given that invariance is built in the notion of topological property (Huneman 2010). Symmetrically, Jones (2014) established that those explanations could adequately explain vulnerability. However, the neutral tunnels in hyperdimensional fitness landscapes are not involved in an explanation of robustness or vulnerability, but rather in explaining adaptive evolution. Hence the category of topological explanations appears more extended than what it initially seemed to be, with respect to its potential targets.

Moreover, considering structural explanations in general, neutral tunnels allow one to refine the understanding of this family of explanations. Adaptation in evolutionary biology mostly traces back to natural selection, whether an adaptation is defined as a result of selection or whether it is merely a possible outcome of natural selection. In turn, selection aims at optimality, which maximizes (inclusive) fitness (e.g., Fisher's fundamental theorem supports this).[6] An adaptationist model figures out what is optimal, hence what selection would reach if acting alone (Potochnik 2009). Then, Rice (2012) argued that optimality explanations are not mechanisms and that they rather pertain to structural explanations, the mathematics involved being differential analysis. What has been shown by the present case of neutral tunnels is that an-

6. In principle, or given some conditions: there is a controversy (e.g., Birch 2016), which is wholly irrelevant here for us.

other structural explanation, involving properties from topology, recently comes into play when adaptive evolution should be explained. Moreover, to the extent that in the case of GP maps neutral networks allow one to explain evolvability, which is the potential for adaptive evolution and therefore constitutes a sort of second-order adaptation,[7] it appears now that, besides optimality explanation, topological explanations constitute a major type of structural explanation biologists use for explaining adaptive evolution. Within the general class of structural explanations, correspondences between topology and differential analysis should thereby inform us about the internal structure of the theory of adaptive evolution.

By uncovering this epistemic configuration, one also casts a light on the epistemic status of recent use of 'neutrality' in evolutionary biology. The emphasis on neutrality is indeed a major feature of population genetics since the 1980s and of ecology in the last decade, to the point that some textbooks consider neutrality as a baseline for the theory; and the meaning of neutrality in the latter decade extended from neutral species or alleles to neutral 'spaces' in hyperdimensional spaces. Thus, as we saw, this shift goes with an extension of the range of structural explanations in evolutionary biology and, especially, the extension of the potential explananda of these explanations—now including adaptive evolution.

While Huneman (2018) theorized the nature of explanatoriness in structural explanations, nothing was said of the kind of issues those explanations were akin to solve. The present investigation shows that neutrality constitutes a state of affairs that structural—and mostly topological—explanations are often fit to handle, since neutrality, featuring no mechanism and difference making, will often require some structural (mathematical) properties to provide explanatory resources, as seen in the cases considered here.

7. Conclusion. Many explanations involving neutral theories and neutral manifolds are aptly interpreted as topological explanations. Thus interpreted, we find that topological explanations are pervasive in evolutionary biology, and we are led to diversify our catalogue of structural and topological explanations in science.

Yet the mathematical objects here at stake—called 'neutral manifolds' in section 5: neutral tunnels and neutral subnetworks—are not together part of a same explanatory project; and even though they should be articulated at some point since they are part of distinct topological perspectives on a same evolutionary reality, the explanatory potential of each of them is quite independent.

---

7. I use 'second-order' in a sense analogous to the concept of 'second-order selection' used to make sense of the selection for properties that increase the potential for selection, such as mutation rate.

REFERENCES

Birch, J. 2016. "Natural Selection and the Maximization of Fitness." *Biological Review* 91 (3): 711–23.

Coyne, R., N. H. Barton, and M. Turelli. 1997. "Perspective: A Critique of Sewall Wright's Shifting Balance Theory of Evolution." *Evolution* 51:643–71.

Craver, C., and L. Darden. 2013. *In Search for Mechanisms: Discovery across the Life Sciences*. Chicago: University of Chicago Press.

Crow, J. 1987. "Neutral Models of Molecular Evolution." In *Neutral Models in Biology*, ed. M. Nitecki and A. Hoffman, 11–25. New York: Oxford University Press.

Frank, S. 2012. "Wright's Adaptive Landscape versus Fisher's Fundamental Theorem." In *The Adaptive Landscape in Evolutionary Biology*, ed. E. I. Svensson and R. Calsbeek, 41–58. New York: Oxford University Press.

Gavrilets, S. 1997. "Evolution and Speciation on Holey Adaptive Landscapes." *Trends in Ecology and Evolution* 12:307–12.

———. 1999. "A Dynamical Theory of Speciation on Holey Adaptive Landscapes." *American Naturalist* 154:1–22.

Gillespie, J. H. 2004. *Population Genetics: A Concise Guide*. Baltimore: Johns Hopkins University Press.

Gross, J. L., and T. W. Tucker. 1987. *Topological Graph Theory*. Reading, MA: Wiley.

Hubbell, S. P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, NJ: Princeton University Press.

Huneman, P. 2010. "Topological Explanations and Robustness in Biological Sciences." *Synthese* 177 (2): 213–45.

———. 2018. "Outlines of a Theory of Structural Explanations." *Philosophical Studies* 175 (3): 665–702.

Jones, N. 2014. "Bowtie Structures, Pathway Diagrams, and Topological Explanation." *Erkenntnis* 79:1135–55.

Kimura, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press.

Lange, M. 2013. "Really Statistical Explanations and Genetic Drift." *Philosophy of Science* 80 (2): 169–88.

Lewis, D. 1973. "Causation." *Journal of Philosophy* 70:556–67.

Lewontin, R. 1974. *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Matthen, M., and A. Ariew. 2002. "Two Ways of Thinking about Natural Selection." *Journal of Philosophy* 49 (2): 55–83.

McGill, B. J., B. A. Maurer, and M. D. Weiser. 2006. "Empirical Evaluation of Neutral Theory." *Ecology* 87:1411–23.

Millstein, R. 2002. "Are Random Drift and Natural Selection Conceptually Distinct?" *Biology and Philosophy* 17 (1): 33–53.

Plutynski, A. 2007. "Drift: A Historical and Conceptual Overview." *Biological Theory* 2 (2): 156–67.

Potochnik, A. 2009. "Optimality Modeling in a Suboptimal World." *Biology and Philosophy* 24 (2): 183–97.

Price, G. R. 1972. "Fisher's Fundamental Theorem Made Clear." *Annals of Human Genetics* 36:129–40.

Rice, C. 2012. "Optimality Explanations: A Plea for an Alternative Approach." *Biology and Philosophy* 27 (5): 685–703.

Stadler, B., P. Stadler, G. Wagner, and W. Fontana. 2001. "The Topology of the Possible: Formal Spaces Underlying Patterns of Evolutionary Change." *Journal of Theoretical Biology* 213 (2): 241–74.

Van Nimwegen, E., J. Crutchfield, and M. Huynen. 1999. "Neutral Evolution of Mutational Robustness." *Proceedings of the National Academy of Sciences* 96 (17): 9716–20.

Winther, R. 2018. *When Maps Become the World*. Chicago: University of Chicago Press.

Woodward, J. 2003. *Making Things Happen*. New York: Oxford University Press.

Wright, S. 1931. "Evolution in Mendelian Populations." *Genetics* 16:97–159.